Desk

Volume 40 May 1954 Number 5

UNIVERSITY OF MICHIGAN
MAY 10 1954
PERIODICAL READING ROOM

Proceedings of the National Academy of Sciences *of the United States of America*

The Proceedings of the National Academy of Sciences

OFFICERS OF THE ACADEMY

DETLEV W. BRONK
President

GEORGE W. CORNER
Vice President

ROGER ADAMS
Foreign Secretary

ALEXANDER WETMORE
Home Secretary

WILLIAM J. ROBBINS
Treasurer

EDITORIAL BOARD OF THE PROCEEDINGS

LINUS PAULING
Chairman

ALEXANDER WETMORE, *Home Secretary*
EDWIN B. WILSON, *Managing Editor*
ROGER ADAMS, *Foreign Secretary*
WILLIAM W. RUBEY, *Chairman of the National Research Council*

J. W. BEAMS
R. E. CLELAND
W. W. RUBEY
HARLOW SHAPLEY
F. E. TERMAN
G. W. BEADLE
E. A. DOISY
G. C. EVANS
C. H. GRAHAM
J. T. PATTERSON
GREGORY BREIT
L. C. DUNN
W. F. LIBBY
J. VON NEUMANN
PAUL WEISS

Subscription price is $7.50 per volume (12 issues); the price of a single issue is $.75 Make all remittances payable to The University of Chicago Press in U.S. currency or its equivalent. Authorized agent for the British Commonwealth, except North America and Australasia: Cambridge University Press, Bentley House, 200 Euston Road, London, N. W. 1, England.

Editorial correspondence should be addressed to the PROCEEDINGS of the NATIONAL ACADEMY OF SCIENCES, Attention Miss Mary D. Alexander, University of Chicago Press, 5750 Ellis Avenue, Chicago 37, Illinois.

Business correspondence should be addressed to the Journals Division, University of Chicago Press, 5750 Ellis Avenue, Chicago 37, Illinois.

Subscribers are requested to notify The University of Chicago Press and their local postmaster immediately of change of address.

Microfilms of complete volumes of this journal are available to regular subscribers only and may be obtained at the end of the volume year from University Microfilms, 313 N. First Street, Ann Arbor, Michigan.

Entered as second-class matter June 21, 1920, at the Post-office at Easton, Pennsylvania, under the Act of August 24, 1912. Accepted for mailing at a special rate of postage provided for in Section 1103, Act of October 3, 1917. Authorized June 18, 1920.

PRINTED IN THE U. S. A.

THE PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES
is published for
THE NATIONAL ACADEMY OF SCIENCES
by
THE UNIVERSITY OF CHICAGO PRESS

Proceedings of the
NATIONAL ACADEMY OF SCIENCES

Volume 40 • Number 5 • May 15, 1954

NUCLEIC ACID PRECURSORS AND PROTEIN SYNTHESIS

By Arthur B. Pardee

VIRUS LABORATORY, UNIVERSITY OF CALIFORNIA, BERKELEY

Communicated by W. M. Stanley, March 19, 1954

It has been suggested that nucleic acids are involved in protein synthesis, and considerable data have been collected, some supporting this view and some conflicting with it.[1] A useful system for the study of this relationship would appear to be enzyme induction in microörganisms, since it has recently been shown that this process involves formation of new protein from amino acids.[2, 3] The determination of enzyme activity as a means of following the formation of small amounts of new protein is advantageous because it is a more sensitive criterion than direct protein analysis.

If a process is essential for protein synthesis, as soon as it is halted enzyme induction should stop. This is in contrast to processes essential for other phases of cell duplication or protoplasmic increase which eventually halt all cell metabolism. Thus, if formation of nucleic acids, rather than simply the presence of nucleic acids, is directly essential for protein synthesis, any process which prevents nucleic acid synthesis should *at the same time* stop enzyme induction. An effect that seemed to be of this sort has recently been observed by Monod *et al.*[3] In testing the requirement for the presence of a number of amino acids or amino acid components for enzyme induction, they also found that a uracil-requiring mutant of *Escherichia coli* was unable to produce β-galactosidase in the absence of uracil.

In the present work, experiments are conducted in which it is shown that removal of the nucleic acid components, purine, pyrimidine, or phosphate, abruptly prevents enzyme induction in *E. coli* strains which require these compounds for growth, and some hypotheses regarding these observations are tested.

Methods.—Various strains of *E. coli* were used in these experiments. Strain 550-460 (requiring cytosine or uracil) was obtained from the American Type Culture collection (ATCC No. 11548). Strains 6386 (requiring pyrimidines) and M45B4 (requiring adenine, guanine, or xanthine), isolated by Dr. B. D. Davis, were obtained from Dr. I. J. Slotnick. Strain B96, isolated by Dr. J. S. Gots, was obtained from Dr. R. Guthrie. Strain M4834 (PABA-less) was isolated by Dr. B. D. Davis.

Cells were grown with aeration by shaking at 37° in a minimal medium[4] [7 gm. K_2HPO_4, 2 gm. KH_2PO_4, 0.5 gm. Na_3-citrate·$5H_2O$, 0.1 gm. $MgSO_4 \cdot 7H_2O$, 1 gm. $(NH_4)_2SO_4$ per liter plus 0.5 per cent glycerol (rather than glucose) as a carbon and energy source] supplemented with the compound necessary for growth. The division time was about 70 minutes. The cells were grown to approximately 7×10^8 cells per

milliliter, washed, and resuspended in a fresh warm sample of the same medium, but containing various amounts of the growth factor. To test the requirement for phosphate, 0.1 *M* trishydroxymethylaminomethane buffer and 0.6 per cent KCl were added in place of phosphate. Excess lactose (0.5 per cent) was added either at once or after 15–30 minutes' aeration by shaking, to induce β-galactosidase for-

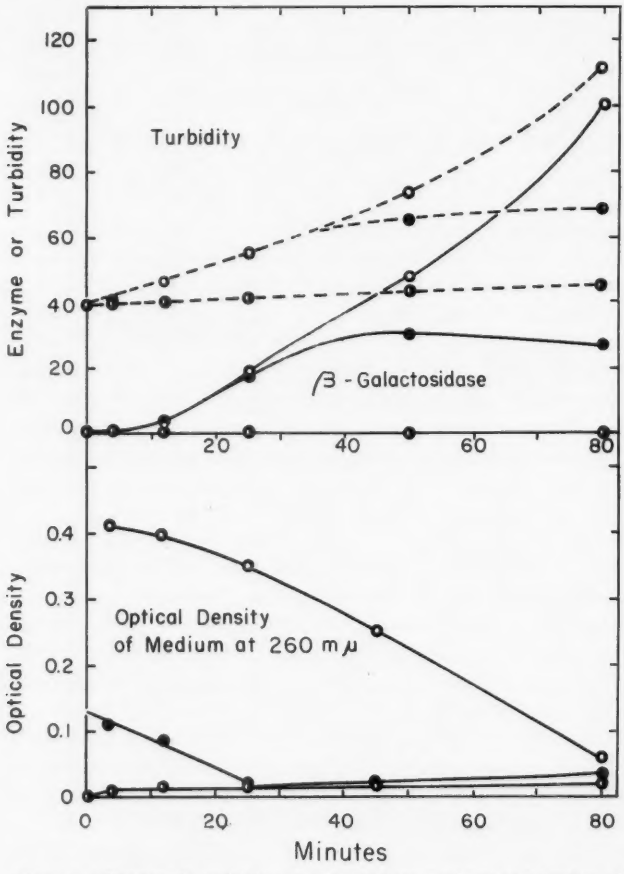

FIG. 1.—Changes in turbidity, β-galactosidase, and optical density of medium at 260 mμ during induction to lactose. ○ = Excess uracil (22 μg/ml); ● = Limiting uracil (2.2 μg/ml); ◑ = No uracil. The optical density with excess uracil is plotted one unit lower.

mation. At various times, samples were removed and added to chloromycetin (15 μg/ml final concentration) and chilled in order to stop further induction.[5] These samples were analyzed to determine changes in enzyme activity and other properties.

The enzyme was assayed by the method of Koppel *et al.*[6] by determination of the amount of *o*-nitrophenyl-β-D-galactoside (NPG) hydrolyzed. The activity was

proportional to time and amount of extract for the extent of reaction studied, so 15 minutes' incubation was used routinely, and results are expressed as micrograms of NPG hydrolyzed per minute per milliliter of original cell suspension.

Turbidity was measured on the Klett-Summerson colorimeter with a green filter and is expressed as Klett reading. For cells grown in the above medium, the proportionality between cells per milliliter and Klett reading is approximately equal to 10^7.

Protein was determined on washed cells by the Folin method.[6] Ribose was determined by the orcinol method[7] and deoxyribose by the method of Ceriotti.[8]

Results.—In Figure 1 are shown typical changes in properties of the pyrimidineless mutant 6386 induced to form β-galactosidase by lactose, at various levels of uracil. It can be seen that enzyme activity and turbidity increased to an appreciable extent only in the presence of uracil. When uracil disappeared from the medium (as measured by light absorption at 260 mμ after centrifuging off the cells), neither enzyme nor turbidity increased further. The enzyme already present was quite stable.

TABLE 1

β-GALACTOSIDASE ACTIVITIES AND TURBIDITIES IN VARIOUS DEFICIENT MEDIUMS

		Strain			
Time (Min.)	Factor*	550-460 (Uracil)†	M45B4 (Adenine)	B96 (Adenine)†	B (Phosphate)
			μg NPG Split/Min/Ml Cells‡		
0	−	2	0		0
4	−	2	0	0	0
12	−	2	0	0.5	0
25	−	2	1	1.5	1
50	−	2	1	7	5
80	−	2	1.5	9	8
25	+	29	22	17	13
80	+	195	73	122	54
			Turbidities in Various Deficient Mediums§		
0		34	29	45	28
80	−	37	29	65	31
80	+	125	64	138	54

* Mediums plus (+) or minus (−) essential factors.
† Medium contained amino acids.
‡ Enzyme activities present at various times after induction.
§ Turbidities expressed as Klett readings.

Almost identical results in all respects were obtained with another pyrimidineless mutant, 550-460, with the purineless mutants M45B4 and B96 and with wild-type *E. coli B* in the absence of phosphate (Table 1). In all cases, induction ceased within 5 or 10 minutes of disappearance of purine or pyrimidine from the medium. The very small production of enzyme in the absence of phosphate may be attributed to release of minute amounts of intracellular phosphate, since 0.17 μg. of phosphate per milliliter of medium was shown to stimulate enzyme production markedly (and this is only about 3 per cent of the phosphate content of the cells in 1 ml. of medium).

A variety of explanations for the above effects, other than obligatory nucleic acid synthesis, come to mind. Some possibilities are tested below.

It seemed possible that the presence of purine or pyrimidine was necessary for the formation of only β-glactosidase (perhaps as part of the enzyme). Therefore, the effect of uracil deficiency on another induction (glycerol oxidation) was tested.

Samples were collected as before at various times during adaptation to glycerol of Strain 550-460. The medium contained 1.0 per cent potassium lactate, salts, and either 0 or 20 μg/ml of uracil. Glycerol was added to the cells at 0 minute. They were collected at later times, washed, and their ability to oxidize glycerol in the presence of 15 μg/ml of chloromycetin was measured by oxygen uptake on the Warburg apparatus. Results showed that appearance of ability to oxidize glycerol was dependent on uracil (Table 2).

TABLE 2

INDUCTION OF GLYCEROL OXIDATION IN PRESENCE AND ABSENCE OF URACIL BY MUTANT STRAIN 550-460

TIME (MIN.)	μL O_2/ML CELL SUSPENSION*		TURBIDITY†	
	20 μg/Ml Uracil	0 Uracil	+ Uracil	0 Uracil
0	0.4	0.4	32	32
90	8.0	0.9	83	40
150	22.0	1.0	172	42

* Enzyme activities at various times after induction.
† Turbidities expressed as Klett readings.

An experiment similar to the one just described, using the Warburg apparatus, was performed, except that adaptation of Strain 550-460 to lactose in the presence of glycerol was measured. The primary purpose of this experiment was to check the NPG assay by an independent method. The results of the two assays were in good agreement. By both methods there was almost no change in activity in absence of uracil, a brief increase with limiting amounts of uracil, and a rapid increase with uracil.

It seemed of interest to see whether uracil, phosphate, and adenine were all necessary for production of the "constitutive" enzyme system oxidizing glucose. This was tested by determining whether ability of cells to oxidize glucose increased in the absence of these compounds. Omission of any one of these compounds from the Warburg flask resulted in a constant ($\pm$10 per cent) or slowly diminishing rate of oxygen uptake on glucose by the appropriate strain of *E. coli*. Results were rather variable in phosphate-free medium but showed no increase on the average; perhaps traces of phosphate were responsible.

Similarly, enzyme activities remained unchanged, or decreased slowly, in the absence of purine or pyrimidine for lactose, lactate, or glycerol oxidation in induced cells. However, it was possible that the net result of synthesis and destruction resulted in a constant enzyme level, if inactivation of β-galactosidase was accelerated in the absence of uracil. This was tested by measuring changes in enzyme activity in the absence of both lactose and uracil of cells induced to lactose. During 80 minutes β-galactosidase activity did not change with experimental error, hence it is concluded that uracil has no effect on stability of this enzyme. Similarly, the lactate-oxidizing system was stable in the absence of uracil and lactate for at least 2.5 hours.

Purineless or pyrimidineless *E. coli* shaken in minimal medium plus glycerol are still able to form colonies on agar for at least 80 minutes. They continued to divide, as shown by microscopic counts and plating, both of which indicated approximately a 40 per cent increase in the number of cells in 80 minutes, but a decrease in cell size was observed.

The uracil-requiring mutant 6386 can use uracil, uridine, or uridine-3-phosphate, as measured by turbidity increase and decrease of optical density, at 260 mμ, of the medium.

In several experiments the protein, ribose, and deoxyribose content of cells at various times of induction and at various levels of purine, pyrimidine, or phosphate were determined. In general, little or no increase of ribose, deoxyribose, or protein was observed in the absence of the essential factor, as compared to a rapid increase of these substances in the complete medium or a limited increase with limiting amounts of the base. Two examples with uracil-requiring mutants are given in Table 3. Although these results show some variation, it is concluded that, while cell division continues briefly in these mutants in the absence of base, there is very little protoplasmic increase (RNA, DNA, or protein synthesis).

TABLE 3

PERCENTAGE CHANGES IN PROTEIN AND NUCLEIC ACIDS AFTER INDUCTION*

TIME (MIN.)	PROTEIN	DNA	RNA
		Mutant 559-460	
No pyrimidine:			
3	100	100	100
10	98	94	93
20	102	89	90
30	94	120	103
50	113	105	92
80	100	76	96
Excess uracil:			
30	134	(190)	145
80	213	272	276
		Mutant 6386	
No pyrimidine:			
0	100	100	100
4	100	104	112
12	108	98	102
25	110	101	111
50	120	106	110
Excess uracil:			
25	123	137	155
50	165	161	214

* Figures represent percentage changes over the initial sample in the presence or absence of uracil.

A possible explanation for the role of purines and pyrimidines is that the compounds are necessary for synthesis of amino acids, which are essential for enzyme production. This hypothesis was tested by growing and inducing mutant strains to lactose utilization in the usual way, except that an amino acid mixture (0.2 per cent Difco casamino acids plus 10 μg/ml tryptophan) was present at all times. The results were the same as in previous experiments (Table 1). Almost no increase of turbidity or of β-galactosidase activity was observed during 80 minutes of induction in medium minus uracil or adenine, as contrasted to a threefold increase of turbidity and a very rapid increase of enzyme activity with excess uracil or adenine. Therefore, the bases must be essential for some process other than amino acid synthesis.

The above results do not permit one to say whether synthesis of DNA or of RNA is necessary for adaptation. Such information would require selective blocking of DNA or RNA synthesis. Herriott has shown that DNA synthesis may be

selectively blocked,[9] and Sher and Mallette (in press) have shown that mustard gas [bis(β-Cl-ethyl)sulfide]-treated cells cannot form DNA but can form the adaptive enzyme lysine decarboxylase. We find that β-galactosidase is also formed by such cells and also by cells in which DNA synthesis is strongly inhibited by ultraviolet light or x-rays. In Table 4 are shown results obtained with *E. coli B* pretreated for 10 minutes at 37° with 7×10^{-4} *M* mustard gas in minimal medium plus glycerol, washed, and induced in the usual way. There was a marked effect on DNA synthesis and relatively little effect on other properties for 40 minutes; then other syntheses fell off. These results indicate that net synthesis of DNA is not necessary for induction.

Cohen and Barner have very recently shown that a thymine-requiring mutant of *E. coli* can adapt in the absence of thymine.[10]

To determine whether other compounds needed for growth (as measured in the usual way, by appearance of turbidity from a very small inoculum within several days) had an effect similar to that of purines or pyrimidines, Mg^{++} was omitted from the medium in one experiment. In another experiment a paraminobenzoic acid (PABA)–requiring mutant, M4834, was placed in a PABA-less medium containing 10 μg/ml sulfanilamide. In the first case, there was inhibition of induction and growth after about 30 minutes. In the second, there was no effect for at least 90 minutes. Therefore, the same effect is not obtained with all "growth" factors, although, as pointed out by Bonner,[11] reserves could permit continued metabolism for some time.

TABLE 4

CHANGES IN MUSTARD GAS–TREATED *E. Coli B* CELLS UPON INDUCTION TO LACTOSE*

	Turbidity	RNA	DNA	Protein	NPG at 40 Min.
Control	156	155	150	145	39
Mustard-treated	136	142	95	139	25

* Figures are percentages of initial values of turbidity, RNA, DNA, or protein, and NPG values (0 at 0 min.) at 40 min. after induction. Fewer than 0.1 per cent cells were viable after mustard gas treatment.

Discussion.—The above experiments show a strict dependence for synthesis of β-galactosidase on the presence of uracil, adenine, and phosphate. These compounds disappear from the medium and are used for nucleic acid synthesis.[12] Such a result suggests that enzyme synthesis requires simultaneous synthesis of one or both types of nucleic acid. Formation of β-galactosidase without DNA synthesis in mustard gas–treated cells indicates that RNA synthesis is more likely to be required.

It may be argued that the cell is in such delicate balance, and all processes are so closely interconnected, that halting any one synthesis stops everything. While this is no doubt true, eventually, there are a number of examples of continued synthesis of one type of material when formation of another is blocked. Mustard gas,[9] ultraviolet light,[13] and lysogenic bacteriophage[14] all block DNA synthesis, with only small effects on RNA or protein synthesis for at least one division time. Also, lysogenic phage and mustard gas do not halt enzyme induction. Jeener and Jeener[15] showed that *Thermobacterium acidophilus R 26*, which requires both uracil and thymidine for extensive growth and cell division, increased in size in the absence of the latter compound, but not when uracil was omitted from the medium.

Another example is found in the work of Gale and Folkes with *Staphylococcus aureus*.[16] They showed that chloromycetin halts protein synthesis without inhibiting total nucleic acid synthesis for some time. All these results are in contrast to the sharp inhibition (within about 5 minutes) of induction in mutants upon exhaustion of purine or pyrimidine from the medium, and it is felt that this is evidence for a close connection between the presence of these compounds and protein synthesis. A similar assumption was made when it was concluded that adaptive enzymes were formed from amino acids, from observed dependence of formation of the former on availability of the latter.[2, 3]

The observation that increase in enzymes needed to oxidize glucose, lactose, glycerol, or lactate, as well as increase of total protein and turbidity (largely increase of protein and nucleic acids), are all closely dependent on the presence of adenine, uracil, and phosphate indicates that protein synthesis in general depends on these compounds. The large amounts of nucleic acids already present in the cell are not useful, either for enzyme synthesis or as a source of small amounts of purines, pyrimidines, or phosphate. One could visualize an active form of RNA perhaps a phosphorylated form like that suggested by Dounce,[17] or perhaps combined with the inducer, to form the organizer,[18] which, once used to mold a specific protein molecule, becomes an inert waste product in *E. coli*. On this basis, one could expect a considerably higher content of RNA in bacteria than in cells which have a high rate of RNA turnover, and this is indeed found. This view is complementary to the one recently expressed by Manson.[19] He showed that there is no turnover of carbon in *E. coli* RNA and suggested that RNA, once formed, is inert in exchanging all its components with other cell materials. Bacterial ribonuclease must be inactive *in vivo*.[19, 20] It is of interest in this regard that no biological activity of isolated RNA or RNA nucleoprotein has yet been clearly demonstrated, except for certain viruses.

There are, of course, other explanations for the role of nucleic acid constituents in protein synthesis, as there must be for so complex a system. For example, these compounds are found in coenzymes and probably are all necessary for protein synthesis in this capacity. However, coenzymes do not soon disappear from the cell in experiments like those described above. Depriving the bacteria of uracil, adenine, or phosphate did not decrease oxidative activity on lactose, glycerol, lactate, or glucose; therefore, oxidative coenzymes are not lost rapidly enough to become limiting. Also, it is stated that in mutants that require vitamins or vitamin precursors, induction does not cease upon removal of the growth factor, in contrast to the present experiments.[11]

In regard to virus multiplication, the above results are of interest in two respects. In bacteriophage-infected *E. coli*, it has been stated that there is no turnover or further formation of RNA.[19, 21] However, the recent result of Hershey[22] indicates some phosphate turnover in RNA, in contrast to the papers cited above. Enzyme induction is blocked,[23] but, since viral protein must be synthesized, though in relatively small amounts, one would postulate that some RNA metabolism must occur. Second, if RNA is inert, it is not clear what its role would be in plant viruses, which contain RNA and protein exclusively. Two possibilities are that protein is the component of these viruses that transmits virus activity or that the virus contains a metabolically active form of RNA.

Summary.—It is shown that strains of *E. coli* which require purines, pyrimidines, or phosphate for growth also require these compounds for induction of enzymes. Within a few minutes of exhaustion of the compounds from the medium, induction of enzymes ceases. The effect is not due to blocking of amino acid synthesis or to destruction of existing enzyme. By inhibition of DNA synthesis with mustard gas, it is shown that DNA synthesis is not necessary for induction. It is suggested that continuous formation of RNA is essential for protein formation and that the RNA found in the bacterial cell is largely an inert metabolic product.

The author is indebted to Dr. D. L. MacDonald and Mr. K. Lonberg for a supply of NPG, to various donors for mutants, and, especially, to Miss Louise S. Prestidge for excellent assistance. This work was aided by a grant from Lederle Laboratories Division of the American Cyanamid Company.

[1] H. Chantrenne, *Symposium sur la biogenèse des protéines: II^e Congrès international de Biochimie* (Sedes, Paris, 1952).

[2] H. O. Halvorson and S. Spiegelman, *J. Bact.*, **64**, 207 (1952).

[3] J. Monod, A. M. Pappenheimer, Jr., and G. Cohen-Bazire, *Biochim. biophys. acta*, **9**, 648 (1952).

[4] J. Lederberg, *Methods in Medical Research*, **3**, 5 (1950).

[5] J. L. Koppel, C. J. Porter, and B. F. Crocker, *J. Gen. Physiol.*, **36**, 703 (1953).

[6] O. H. Lowry, N. J. Rosebrough, A. L. Farr, and R. J. Randall, *J. Biol. Chem.*, **193**, 265 (1951).

[7] W. C. Schneider, *J. Biol. Chem.*, **161**, 293 (1945).

[8] G. Ceriotti, *J. Biol. Chem.*, **198**, 297 (1952).

[9] R. M. Herriott, *J. Gen. Physiol.*, **34**, 761 (1951).

[10] S. S. Cohen and H. D. Berner *Fed. Proc.*, **13**, 193 (1954).

[11] D. M. Bonner, in W. D. McElroy and B. Glass (eds.), *Phosphorus Metabolism*, **2**, 153 (1952).

[12] A. L. Koch, F. W. Putnam, and E. A. Evans, Jr., *J. Biol. Chem.*, **197**, 105 (1952).

[13] A. Kelner, *J. Bact.*, **65**, 252 (1953).

[14] L. Siminovich, *Ann. Inst. Pasteur*, **84**, 265 (1953).

[15] H. Jeener and R. Jeener, *Exp. Cell Res.*, **3**, 675 (1952).

[16] E. F. Gale and J. P. Folkes, *Biochem. J.*, **53**, 483 (1953).

[17] A. L. Dounce, *Enzymologia*, **15**, 251 (1952).

[18] M. R. Pollock, in E. F. Gale and R. Davies (eds.), *Adaptation in Microorganisms* (Cambridge University Press, 1953).

[19] L. A. Manson, *J. Bact.*, **66**, 703 (1953).

[20] A. B. Pardee and R. E. Kunkee, *J. Biol. Chem.*, **199**, 9 (1952).

[21] S. S. Cohen, *J. Biol. Chem.*, **174**, 295 (1948); L. M. Kozloff, K. Knowlton, F. W. Putnam, and E. A. Evans, Jr., *J. Biol. Chem.*, **188**, 101 (1951).

[22] A. D. Hershey, *J. Gen. Physiol.*, **37**, 1 (1953).

[23] J. Monod and E. Wollman, *Ann. Inst. Pasteur*, **73**, 937 (1947).

*SYNTHESIS OF AROMATIC COMPOUNDS BY NEUROSPORA**

By E. L. Tatum, S. R. Gross,† G. Ehrensvärd,‡ and L. Garnjobst

DEPARTMENT OF BIOLOGICAL SCIENCES, STANFORD UNIVERSITY

Communicated by E. L. Tatum, March 19, 1954

Increasing insight into the biosynthesis of aromatic substances has been obtained during the past few years primarily through the study of mutant strains of microorganisms,[1] and by the use of isotopic carbon.[2] Shikimic acid, reported to be involved in the biosynthesis of the aromatic amino acids and of *p*-aminobenzoic acid by *Neurospora*[3] and *Escherichia coli*,[4] has been shown to be derived in *E. coli* from dehydroquinic acid by way of dehydroshikimic acid.[5] This paper reports further genetic and biochemical studies of a mutant of *Neurospora crassa* which requires for growth phenylalanine, tyrosine, tryptophan, and *p*-aminobenzoic acid (PAB).

The mutant strain Y7655a (*arom*) was isolated following treatment of *Neurospora* macroconidia with methyl-bis(β-chloroethyl)amine[6] and its multiple aromatic requirements identified by S. E. Reaume. Detailed genetic analysis based on examination of 145 asci from five crosses[7] has shown the *arom* locus to be in the right arm of linkage group II, distal to *pe* and proximal to *ac*, with an uncorrected centromere distance of 21.7. The multiple requirement was retained in all cultures isolated from these various crosses. There is some evidence, however, that the PAB requirement may be altered by modifying genes. The qualitative and quantitative growth responses reported here were obtained with the *arom* strain (Y7655a), and a double mutant *arom, tryp-1* (Y7655, 10575), was used for the isotope experiment.

The quantitative requirements of *arom* for PAB, tryptophan, phenylalanine, and tyrosine are of the same order of magnitude as those of other mutant strains deficient only in the single syntheses. A large number of compounds have been examined auxanographically for their ability to substitute for the aromatic substances, either singly or collectively. The results, given in Table 1, show the specificity of the requirements of this mutant strain. Only one compound, shikimic acid, is by itself able to support growth in minimal medium, although rather large amounts are necessary for optimal growth (5 mg/20 ml). These results suggest that the *arom* locus is concerned with the production of shikimic acid, which serves as the common precursor of PAB, tryptophan, phenylalanine, and tyrosine. This concept is supported by the finding that, in the presence of any three of these substances, shikimic acid supports growth in quantities proportional to the amount of the omitted compound normally required. Also in accord with this interpretation are the activities of anthranilic acid and indole in replacing tryptophan. Kynurenine probably owes its activity to its conversion to anthranilic acid.[8] Shikimic acid is least effective in replacing phenylalanine for the *arom* mutant of *Neurospora*, as has also been reported for *E. coli*.[9] The comparative inefficiency of the conversion of shikimic acid to phenylalanine in *Neurospora* appears to be due in part to an inhibitory effect of tryptophan. Added tryptophan (but not indole) inhibits growth on shikimic acid, unless phenylalanine is present; apparently, exogeneous tryptophan inhibits the conversion of shikimic acid to phenylalanine.

Among the inactive compounds listed in Table 1, dehydroquinic acid (DHQ) and dehydroshikimic acid (DHS) have been shown to be precursors of shikimic acid

in *E. coli*.[5] Their inactivity indicates either that this is not true for *Neurospora* or that the *arom* mutant is blocked in the conversion of DHS to shikimic acid. The following experiments were designed to test these hypotheses.

In order to examine culture filtrates of the *arom* strain for the accumulation of compounds involved in aromatic synthesis, use was made of bacterial mutants deficient in the synthesis of shikimic acid. Strains of *E. coli*, 83-1, blocked between DHQ and DHS,[9] and 156-53D2, a double mutant blocked both before and after DHS,[10] as well as a strain of *Aerobacter aerogenes*, A170-143S1, blocked before DHQ,[9] were used for this purpose. In the absence of shikimic acid, these strains differ

TABLE 1

COMPOUNDS TESTED FOR GROWTH PROMOTION OF THE *arom* STRAIN

COMPOUND	COMMENTS
Active compounds:	
Shikimic acid	Substitutes for each or all requirements
Anthranilic acid	Substitute for tryptophan only
Indole	
Kynurenine	
Inactive compounds (tested auxanographically):	
Dehydroquinic acid	Tested for replacement of shikimic acid
Dehydroshikimic acid	
Acetyl-dehydrophenylalanine	Tested for replacement of phenylalanine or tyrosine
Dopa	
p-Hydroxyphenylpyruvic acid	
Indole-acetic acid	Tested for replacement of tryptophan
Nicotinamide	
Tryptamine	
D-Tryptophan	
Aniline	Tested for replacement of each requirement and of multiple requirements
Benzoic acid	
trans-Cinnamic acid	
Crotonic acid	
Cyclohexane-carboxylic acid	
p-Chlorobenzoic acid	
Gentisic acid	
Gentisyl alcohol	
p-Hydroxybenzoic acid	
Inositol	
Kynurenic acid	
α-Phenylglycine	
Pimelic acid	
Protocatechuic acid	
Pyrogallol-carboxylic acid	
Quinic acid	
Salicylic acid	
Sorbic acid	
Xanthurenic acid	

from the *arom* mutant in requiring *p*-hydroxybenzoic acid (POB), in addition to tryptophan, tyrosine, phenylalanine, and PAB. Since the requirements for PAB and POB are small compared to those for the aromatic amino acids, the detection of compounds involved in aromatic synthesis was greatly enhanced by plating washed suspensions of bacteria in the presence of tyrosine, tryptophan, and phenylalanine.

Bioautography of butanol–acetic acid–developed paper chromatograms of culture filtrates of the *arom* strain grown on shikimic acid or on the four aromatic substances revealed the presence of a compound which could satisfy the PAB and POB requirements of strains A170-143S1 and 83-1 but not those of strain 156-53D2. The R_f of this compound was drastically altered by the presence of tartrate in the

culture filtrate applied to the paper. In the absence of tartrate the compound had an R_f identical with that of DHS, slightly greater than that of shikimic acid, and much greater than that of DHQ. In the presence of tartrate the R_f of the accumulated compound was depressed, while that of shikimic acid was unchanged. The R_f of DHS was equivalently depressed when applied to the paper in the presence of tartarate-containing culture filtrates. Acidification with HCl of the culture filtrates containing tartrate restored the R_f values of the unknown and of DHS to the R_f obtained in the absence of tartrate.

Strain 156-53D2 will respond only to those compounds subsequent to DHS in the pathway of PAB and POB biosynthesis, while DHS can replace the PAB and POB requirements of strains 83-1 and A170-143S1 and DHQ can do so only for strain A170-143S1. Since both A170-143S1 and 83-1 responded to the accumulated compound, its position in the pathway of aromatic synthesis must be between DHQ and shikimic acid. Comparison of the biological specificity and chromatographic mobility of the accumulated compound with those of DHS indicates that the two compounds are identical.

The rate of accumulation of DHS, as determined biologically, is quite slow; 1–3 μgm. per milliliter of culture filtrate is accumulated within 120 hours of growth, and thereafter the concentration decreases slowly. It was noted, however, that culture filtrates showed a UV absorption below 260 μm. much greater than would be expected on the basis of the DHS accumulated. Furthermore, the addition of $FeCl_3$ to the filtrates resulted in the development of a green color characteristic of protocatechuic acid.

Protocatechuic acid (PCA) was isolated from a 15-liter culture filtrate of the *arom* strain grown for 4 days in the presence of 450 mg. of L-tyrosine, 450 mg. of L-phenylalanine, 150 mg. of indole, and 0.6 mg. of PAB. The volume of filtrate was reduced *in vacuo* and the solution extracted continuously with ether for 48 hours. The ether solution was evaporated, and the PCA was crystallized from dry ether-chloroform. The isolated acid was recrystallized three times from ether-chloroform and ethanol-benzene mixtures. The material isolated had an R_f in butanol–acetic acid equal to that of PCA and yielded catechol upon heating. The melting point of the isolated material was 203–205° C., authentic PCA melted at 204–205° C., and the melting point of a mixture was 204–205° C. Large quantities (up to 2 mg/20 ml) of PCA are accumulated in culture filtrates of the *arom* strain during the first 4 days of growth, but PCA is subsequently quantitatively metabolized. PCA does not satisfy the requirement for either PAB, tyrosine, phenyl alanine, or tryptophan of any of the bacterial strains studied.

From growth responses it appears that the *arom* mutation blocks the conversion of DHS to shikimic acid. This view is supported by the demonstrated accumulation of DHS. The small amount of DHS produced by *arom* cultures as compared with that produced by the analogous mutant of *E. coli*[5] may result from the conversion of DHS to PCA. Since this conversion does not take place in *E. coli*, for which DHS is inhibitory,[10] the conversion by *Neurospora* may serve as a detoxification mechanism. The most reasonable source of PCA appears to be either DHS or a closely related compound. The conversion of DHS to PCA is known to take place upon heating.

The conclusions based on growth responses of the *arom* strain and on the accumu-

lation of DHS and PCA were further examined by carrying out an experiment using C^{14}-labeled sucrose. For this experiment the double mutant *arom-tryp-1* (Y7655-10575) was grown on shikimic acid and indole. The *tryp-1* strain is unable to convert anthranilic acid to indole. The double mutant *arom-tryp-1* requires for growth shikimic acid (replaceable by PAB, phenylalanine, and tyrosine) and tryptophan (replaceable only by indole). The *arom-tryp-1* strain accumulates anthranilic acid only in the presence of shikimic acid.

After 4 days of incubation at 30° C., the mycelium and the culture medium were separated by filtration. Anthranilic acid and PCA were isolated from the culture filtrate as follows: The filtrate was acidified to pH 3 and extracted with ether. The ether extract was chromatogramed on paper using butanol–acetic acid. The anthranilic acid and PCA were eluted from appropriate strips cut from the chromatogram. Anthranilic acid was determined fluorimetrically in the eulate, diluted with 100 mg. of anthranilic acid, and the diluted sample was purified by recrystallization and sublimation *in vacuo* to constant activity. The PCA was determined in the eluate colorimetrically by the $FeCl_3$ test. It was then diluted with 100 mg. of authentic PCA, and the product was purified by recrystallization in ethanol-benzene to constant activity. Both materials were decarboxylated by heating in sealed tubes—anthranilic acid for 1 hour at 205° C. and PCA for 1 hour at 250° C. The CO_2 formed from each carboxyl group was released in an evacuated system and trapped as $BaCO_3$. Essentially quantitative yields were obtained in both instances.

Tyrosine and alanine (as a representative aliphatic amino acid) were obtained from the mycelium by hydrolyzing the washed and dried material (32 mg.) in 6 *N* HCl at 120° C. for $^1/_2$ hour. After removal of the excess HCl over KOH, solutions were adjusted to pH 6 with NaOH and chromatogramed on paper with butanol–propionic acid. The tyrosine and alanine areas were cut out and eluted with water. Both samples were rechromatogramed in a two-dimensional system with phenol followed by butanol–propionic acid. The alanine was eluted with water, the concentration was determined by the modified ninhydrin method of Troll and Cannan,[11] and 100 mg. of unlabeled DL-alanine was added. The product was then recrystallized from 90 per cent ethanol to constant activity. The application of the quantitative ninhydrin method to the tyrosine eluate proved unsatisfactory, possibly because of some destruction of tyrosine with phenol or the presence of interfering substances. The dilution factor for tyrosine could not be determined, and the figures given in the table are uncorrected. The eluted tyrosine was diluted with 100 mg. of L-tyrosine, and after treatment with Norite the product was recrystallized from water to constant activity. Alanine and tyrosine were decarboxylated with ninhydrin[12] and the carboxyl carbon recovered as $BaCO_3$.

The results of the isotope analyses given in Table 2 show that protocatechuic acid and alanine are uniformly labeled within acceptable range of the specific activity of the sucrose provided. In contrast, anthranilic acid is relatively unlabeled and is therefore derived primarily from shikimic acid. This view is supported by the observation that the *arom-tryp-1* mutant does not accumulate anthranilic acid when grown in the presence of phenylalanine, tyrosine, and PAB instead of shikimic acid. The addition of shikimic acid to cultures grown under these conditions results in the accumulation of large amounts of anthranilic acid within 4–5 hours.

Since the isolated anthranilic acid is uniformly labeled and has approximately

one-seventh the specific activity of the PCA, it appears that (1) shikimic acid is converted to anthranilic acid and (2), under the conditions used for growth, the *arom* block is not complete. The activity of PCA is consistent with its proposed production from DHS, which the *arom* strain cannot convert to shikimic acid.

The relative activities of the isolated tyrosine and of the tyrosine carboxyl carbon are consistent with the hypothesis that the ring carbons are derived from shikimic acid and the side-chain carbons from other sources. Although in these experiments the specific activity of the undiluted tyrosine is unknown, on the assumption that the side-chain carbons are equally labeled, the relative activities of the side-chain and ring carbons can be calculated as shown in Table 2. This assumption and the proposed derivation of tyrosine are consistent with the finding that the calculated ratio of activities of the carbon atoms of the side chain and ring of tyrosine (6.4) is essentially the same as the ratio between the activities of PCA and of anthranilic acid (7.1).

TABLE 2

CONVERSION OF C^{14}-SUCROSE TO ANTHRANILIC ACID, PROTOCATECHUIC ACID, TYROSINE, AND ALANINE BY MUTANT STRAIN *arom, tryp-1*

(Culture grown 4 days at 30° C. with 100 mg. C^{14}-sucrose, 10 mg. shikimic acid, and 0.5 mg. indole in 20 ml. tartrate-free medium)

COMPOUND	RADIOACTIVITY*: Disintegrations per 0.1 mM Compound	RADIOACTIVITY*: Disintegrations per 0.1 mM Carbon
Sucrose	3.61×10^7	3.01×10^6
Anthranilic acid (ANTH)	2.20×10^6	3.14×10^5
Anthranilic acid carboxyl carbon	...	2.88×10^5
Protocatechuic acid (PCA)	1.56×10^7	2.23×10^6
Protocatechuic acid carboxyl carbon	...	2.47×10^6
Alanine	7.62×10^6	2.54×10^6
Alanine carboxyl carbon	...	2.71×10^6
Tyrosine	2,380†	...
Tyrosine carboxyl carbon	...	605†

$$C^{14} \text{ ratio } \frac{\text{PCA}}{\text{ANTH}} = \frac{2.23 \times 10^6}{3.14 \times 10^5} = 7.1;$$

$$C^{14} \text{ ratio } \frac{\text{Tyrosine side-chain carbon}}{\text{Tyrosine ring carbon}} = \frac{605}{1/6(2{,}380 - 3 \times 605)} = 6.44.$$

* Activity of isolated material after correction for dilution with carrier.
† Activity per 0.1 mM compound of isolated diluted material. Exact dilution unknown.
Samples plated on aluminum planchets and counted in a gas flow counter to within 4 per cent standard error; values corrected for counter efficiency.

Summary.—The aromatic-less (*arom*) mutant strain of *N. crassa* has been found to grow only on a mixture of four aromatic substances, para-aminobenzoic acid, tryptophan, phenylalanine, and tyrosine, or on shikimic acid but not on dehydroquinic or dehydroshikimic acids. The *arom* locus is located on the right arm of linkage group II, 21.7 map units from the centromere.

The *arom* strain has been found to accumulate small amounts of dehydroshikimic acid and relatively large amounts of protocatechuic acid.

It is concluded that the synthesis of aromatic rings in *Neurospora* follows a pathway similar to that in *E. coli*, in that dehydroshikimic acid is the normal precursor of shikimic acid, with the *arom* mutant blocked in this conversion. In *Neurospora*, however, the dehydroshikimic acid accumulated as a consequence of the block is apparently converted to protocatechuic acid.

The results of the isotope experiments reported are in accord with these conclusions and, in addition, indicate that all the carbon atoms of anthranilic acid and of the aromatic nucleus of tyrosine are derived fairly directly from shikimic acid.

The authors are very grateful to Dr. H. O. L. Fisher, of the University of California, and to Dr. Karl Pfister, of Merck and Company, for generous supplies of shikimic acid; and to Dr. B. D. Davis, of the New York University School of Medicine, for mutant cultures of *E. coli* and for samples of DHQ and DHS.

* Supported in part by Research Grant C-2167 from the National Cancer Institute of the National Institutes of Health, United States Public Health Service.

† Public Health Service Postdoctorate Research Fellow of the National Cancer Institute.

‡ On leave from the Wenner-Grens Institut, Stockholm, Sweden, under a research grant from Eli Lilly and Company.

[1] B. D. Davis, *Symposium sur le métabolisme microbien: II^e Congrès international de Biochimie* (1952), p. 32; C. H. Haddox, these PROCEEDINGS, **38,** 482 (1952); M. Gordon, F. A. Haskins, and H. K. Mitchell, these PROCEEDINGS, **36,** 427 (1950).

[2] G. Ehrensvärd and L. Reio, *Ark. f. Kem.*, **5,** 229 (1953).

[3] E. L. Tatum, in F. Skoog (ed.), *Plant Growth Substances* (Madison: University of Wisconsin Press, 1951).

[4] B. D. Davis, *J. Biol. Chem.*, **191,** 315 (1951).

[5] I. I. Salamon and B. D. Davis, *J. Am. Chem. Soc.*, **75,** 5567 (1953).

[6] E. L. Tatum, R. W. Barratt, N. Fries, and D. Bonner, *Am. J. Bot.*, **37,** 38 (1950).

[7] R. W. Barratt, D. Newmeyer, D. D. Perkins and L. Garnjobst, *Advances in Genetics*, Vol. 6 (in press) (1954).

[8] F. A. Haskins and H. K. Mitchell, these PROCEEDINGS, **35,** 500 (1949).

[9] B. D. Davis, *J. Bact.*, **64,** 729 (1952).

[10] B. D. Davis, *ibid.*, p. 749.

[11] W. Troll and R. K. Cannan, *J. Biol. Chem.*, **200,** 803 (1953).

[12] D. D. Van Slyke, R. T. Dillon, R. T. MacFadyen, and P. Hamilton, *J. Biol. Chem.*, **141,** 627 (1941).

EFFECT OF LOCAL-ACTION CURRENTS ON THE IRON POTENTIAL

BY HERBERT H. UHLIG

CORROSION LABORATORY, DEPARTMENT OF METALLURGY, MASSACHUSETTS INSTITUTE OF TECHNOLOGY

Communicated by C. R. Soderberg, November 10, 1953

T. W. Richards and W. T. Richards[1] in 1924 pointed out that the iron electrode potential is appreciably affected by change of hydrogen-ion concentration, even though the ferrous-ion concentration remains constant. This observation was confirmed subsequently by many others, a general summary of the data up to 1938 being provided by Gatty and Spooner.[2] By and large, the iron potential is more active than the hydrogen electrode potential but is truly linear with pH, the slope of which, in acid mediums of pH less than about 5–6, approaches or is less than the theoretical 0.059 volt per pH unit. In less acid or in alkaline mediums, the slope is definitely below the theoretical, averaging perhaps 0.014 volt per pH unit. The data of various investigators, although reasonably concordant in this respect, are usually not in agreement on actual values of potentials, even though measurements

may have been made using refined experimental control of all the obvious variables. Richards, and several other investigators who followed, showed that the steady-state potential of pure iron obtained from various sources is not reproducible to better than a hundredth of a volt or more. One of the purposes of the present paper is to point out the basic reasons for this lack of reproducibility.

Recently, D'Ans and Breckheimer[3] reported several measurements for carbonyl and electrolytic iron in solutions of 1 *M* KCl plus HCl, 1 *M* Na_2SO_4 plus H_2SO_4, 3 *M* $(NH_4)_2SO_4$ plus H_2SO_4, and 1 *M* Na acetate plus acetic acid, with or without addition of ferrous salts, as a function of pH. In their measurements, each of which required several days, they found that within the pH range 1–3 or 1–5, depending on the electrolyte, the slope of potential versus pH was approximately 0.059 and was independent of ferrous-ion concentration. As the pH approached higher values, the potential was no longer dependent on hydrogen ion but responded to ferrous-ion activity instead, approximately in the manner predicted by the Nernst equation. The potential of iron at a given pH was not the same in all electrolytes but was increasingly noble in the order of the solutions listed above. Electrolytic iron agreed with carbonyl iron to no better than 30 mv. Additional data along the same lines were furnished by Bonhoeffer and Jena.[4]

Measurements of Standard Potential.—Despite the inherent lack of reproducibility and uncertain reversibility, values for the standard potential, $E°$, corresponding to the reaction

$$Fe \rightarrow Fe^{++} + 2e$$

have been calculated from potential measurements by several investigators. Early values of $E°$ varied from 0.36 to 0.66 volt, but in 1926 Hampton[5] arrived at the value 0.441 volt, close to the presently accepted value. He employed the following cell: Fe, 0.1 *M* $FeCl_2$, Hg_2Cl_2, Hg, which showed, he said, variations of potential typical of iron and "hard to fully explain." Iron amalgam and reduced iron oxide powder electrodes behaved reasonably well, but annealed iron wires gave values of potential about 0.05 volt lower. Later Randall and Frandsen,[6] using the identical cell and concentration of electrolyte, obtained the value which now appears in tables of standard potentials, namely, 0.440 volt. The fact that the same cell and the same concentration of $FeCl_2$ were used probably explains the good agreement. In addition, the reported measurements of both investigators were taken over a period of weeks or, in some cases, months, which served to establish the same pH (probably near neutral) at the surface of the electrodes aiding the agreement.

Recently, Patrick and Thompson,[7] using carbonyl iron and either ferrous chloride or ferrous sulfate in concentrations slightly more dilute than 0.1 molal, report a value of $E°$ equal to 0.409 volt. They ascribe their lower value to less hydrogen in carbonyl iron compared with electrolytic iron or hydrogen-reduced iron oxide, and to their use of thoroughly deaerated electrolyte. In view of the previously reported potential behavior of iron, it is also possible, of course, that the pH of their electrolyte differed from that of Hampton and of Randall and Frandsen. Richards[1] and D'Ans and Breckheimer[3] also believed that occluded hydrogen accounted for the observed erratic potential behavior of iron, and Gatty and Spooner leaned to the possibility of a surface layer of hydride. Hampton, however, doubted the effect of hydrogen, in view of one experiment which showed that iron powder re-

duced from the oxide by carbon monoxide gave a high initial value of potential similar to that of iron prepared by reduction with hydrogen, and believed instead that variable small particle size of the powder was the cause.

Although potential measurements may approach the equilibrium value of iron in a ferrous salt solution, as is indicated by the reasonable correspondence of most recent values of E°, derived from potential data, with indirect thermodynamic calculations, it has been shown that the true reversible potential of any corrodible metal, in fact, is never achieved.[4, 8, 9] The extent of deviation from reversibility depends on electric currents which polarize discrete anodic and cathodic areas of the metal surface. Equilibrium conditions are disturbed, accounting for an observed potential that may change with time and that is always more cathodic than the thermodynamic potential, in accordance with the relation

$$E_{\text{therm.}} = E_{\text{obs.}} + f(i),$$

where $f(i)$ is a function of local-action current, i.

The discussion which follows presents evidence that local-action currents for iron are so large as to produce serious deviations from reversibility. Also, the variable steady-state potentials are thought to be caused not directly by hydrogen or hydride or particle size but by variable corrosion rates of iron. Interstitial hydrogen, of course, may be one of the factors determining corrosion rate, but other factors are equally or more important, such as trace impurities in the iron, and composition and pH of the electrolyte.

Effect of pH on the Iron Potential.—Whether iron reacts with acid to form hydrogen gas and a corresponding ferrous salt or reacts with water at a higher pH to form ferrous hydroxide and hydrogen, the free-energy change for the reaction is negative, and the reaction, therefore, is able to and does proceed. The free-energy change for reaction with water, for example, is

$$\text{Fe}(s) + 2\text{H}_2\text{O}(l) \rightarrow \text{Fe(OH)}_2(s) + \text{H}_2(g), \quad \Delta F^\circ = -2{,}300 \text{ cal.}$$

The rate of this reaction in deaerated solutions is controlled by the rate at which hydrogen is evolved at cathodic areas of the metal surface. Since hydrogen overvoltage is a measure of this rate at a given applied potential, the hydrogen overvoltage of cathodic areas is also a measure of the corrosion rate of iron. This is true whenever the metal reaction is accompanied by hydrogen evolution and in the absence of depolarizers. The local-action or corrosion currents accompanying slow chemical attack polarize the anode potential to a more cathodic value simultaneously with the polarization of the cathode potential to a more anodic value. The value of current corresponding to the intersection of the cathodic and anodic polarization curves represents the maximum value that can be attained by the short-circuited cells and, hence, is called the "maximum corrosion current." The common potential achieved by both polarized electrodes is called the "corrosion potential," which, it will be recognized, is the steady-state or compromise potential of the composite of anodes and cathodes on the metal surface. This situation is illustrated in Figure 1. The iR drop, where R is the resistance of the electrolyte and metal paths between anodes and cathodes, can be considered negligible. This, in fact, is usually realized for most metals, since anodes and cathodes are relatively near each other. Strictly speaking, the anode and cathode polarization curves

never intersect but closely approach each other to within a value of potential represented by the sum of small iR drops through the surface electronic and ionic circuits. The term E_a is the truly reversible (to Fe^{++} only) or zero-current potential for iron, and E_c is the cathodic potential, which, in the case of electrode reactions involving evolution of hydrogen, is the equilibrium hydrogen electrode potential at 1 atm. If the pH of the environment is altered from pH_1 to the more acid value pH_2, then E_c moves in the more noble direction to the extent of 0.0592 ($pH_1 - pH_2$) volt. The resultant corrosion current increases, and the corrosion potential is simultaneously displaced in the cathodic direction. Since observed potentials of iron *are always the corrosion potential* corresponding to E_1 and are not the equilibrium potential E_a, the dependence of the observed potential on pH is readily understood. The actual deviation of E_1 from the thermodynamic potential E_a depends on the corrosion rate measured by i and the extent to which the anodic areas of the metal are polarized by i. Since the corrosion rate, in turn, depends on the source of iron, the nature of the anions present, and the concentration of dissolved oxygen, as well as on the pH, the corrosion potential or observed potential is subject, consequently, to all the variations that affect the corrosion behavior. This is true of potentials measured in various salt solutions where Fe^{++} activity reaches an unknown but relatively constant value at the metal surface and is true as well of the refined measurements in ferrous salt solutions from which E° values have been calculated. It is also obvious from Figure 1 that, by reducing the corrosion rate through the use of neutral or alkaline mediums and by excluding oxygen, which, as a depolarizer, increases local-action currents, the value of corrosion potential, E_1, can be made to approach the reversible potential, E_a. This situation, in effect, accounts for what must now be considered a fortuitous approximation of recent E° values for iron to values calculated indirectly from thermodynamic data.

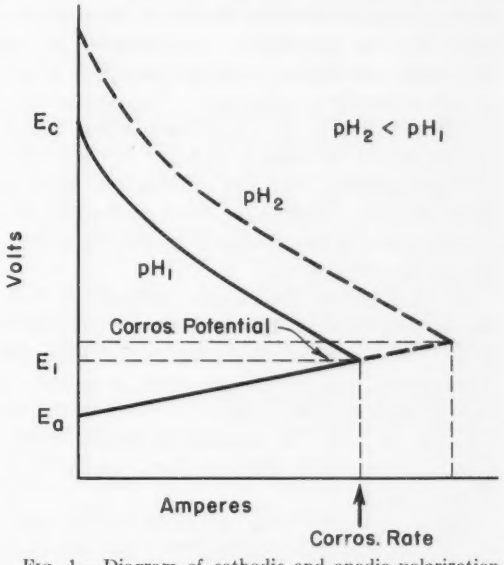

FIG. 1.—Diagram of cathodic and anodic polarization curves for a metal in a solution at two values of pH.

Dependence of the Corrosion Potential on Corrosion Rate.—The value of E_c corresponding to the hydrogen electrode depends on pH only. The cathodic polarization behavior, in turn, is a function of two terms. The first is a hydrogen overvoltage term, $\beta \log i/i_0$, where β is a constant usually expressed as $2.303RT/\alpha F$ and i is the current density corresponding to local-action currents. The exchange current, i_0, corresponds to the equilibrium reaction $H_2 \leftrightarrows 2H^+ + 2e$. The second

term is a concentration-polarization term equal to (RT/F) ln $[i_d/(i_d - i)]$, where i_d is the limiting current density.[10] Accordingly, the polarized cathodic potential E_c', assuming a potential change in the cathodic direction as positive, is equal to

$$E_c' = -0.0592 \text{ pH} - \beta \log \frac{i}{i_0} - 2.303\, RT/F \log \frac{i}{i_d - i} \qquad (1)$$

Less is known about the correct polarization function for the anode, especially since measurements have invariably reported change of the corrosion potential with applied polarizing current rather than for the anode alone in the absence of local-action currents. Difficulties of measurement apart from this important factor have been summarized by Butler,[11] who pointed out that the true surface area of the anode may change continuously during anodic dissolution and also that the ohmic potentials between the tip of the reference electrode and the measured electrode may be appreciable. Furthermore, an additional consideration should be mentioned, relating to a possible change in ratio of surface anode to cathode areas as the electrode is polarized. Changes of this kind probably account for the marked hysteresis of anodic polarization employing increasing applied current compared with decreasing current.[4, 12] The fact that hysteresis is less or is absent for corresponding cathodic polarization measurements suggests that in at least some aqueous mediums the total anodic area of iron is much less (hence the percentage effect of any change is greater) than the cathodic area. Later calculations support this suggestion.

We shall assume that anodic polarization based on true anodic area is a linear function of current density. Anodic polarization of cadmium[9] and zinc amalgams,[13] for example, where the local-action currents are extremely small, follow this relation, and it is thought that, in general, all metals polarize linearly at sufficiently small applied currents.[14] Note should be taken that the transition metals, e.g., Fe, Ni, and Co, polarize anodically much more than do the nontransition metals, a fact which becomes significant in later comparisons of calculation with observation.

Hence, following what is considered the best assumption at the present time, the polarized anodic potential, E_a', can be expressed as

$$E_a' = -E_a + \frac{bi}{A} + 2.303 \frac{RT}{2F} \log \frac{i + k(Fe^{++})}{k(\text{Fe}^{++})}, \qquad (2)$$

where E_a conforms to the Lewis and Randall sign convention, b is a constant based on total unit area, A is the fraction of total area represented by the anode, and the log term expresses anodic concentration polarization as a function of i, a constant k, and the activity of ferrous ion in the bulk of the solution.[15] The value of A increases to a maximum value of unity with increasing externally applied current used to polarize the electrode anodically, but it is not expected to change markedly in the absence of external polarization within the usual range of pH or ferrous-ion concentration. For the experimental conditions considered at present, therefore, A will be assumed a constant.

The usual measurement of potentials implies the necessity of very small rates of attack by the electrolyte; consequently, the situation in which we are primarily interested is for relatively low current density i, or for corrosion rates so small that

concentration-polarization terms can be neglected as a first approximation. Next, it is apparent that, since E_a' and E_c' become identical at the corrosion potential E_1 corresponding to maximum local-action current i, the value of i from equation (2) is

$$i = \frac{(E_1 + E_a)A}{b}, \tag{3}$$

and, from equation (1),

$$E_1 = -0.0592 \text{ pH} - \beta \log \frac{(E_1 + E_a)A}{b} + \beta \log i_0. \tag{4}$$

Differentiating with respect to pH, on the assumption that the equilibrium potential, E_a, and the hydrogen overvoltage are independent of pH, or β, E_a, b, and i_0 are constants,

$$\frac{dE_1}{d\text{pH}} = \frac{-0.0592}{1 + (\beta A/2.303bi)}. \tag{5}$$

This equation shows that at low values of local-action current i corresponding to a small corrosion rate, dE_1/dpH, is small, and, accordingly, the observed potential is not sensitive to change in pH. On the other hand, at larger values of i the slope of potential versus pH approaches the value -0.0592. Recent measurements of Stern[16] show that the corrosion rate of iron in deaerated 4 per cent NaCl of pH 0.96 is 29 mg/dm^2/day (mdd), 29 mdd at pH 2.9, and 25 mdd at pH 4. Hence i is practically constant within this range of pH where the corrosion rate is relatively high, and, therefore, E_1 should be linear with pH, as has been previously observed by many investigators. Furthermore, equation (5) indicates that the absolute value of slope is always less than 0.0592.

In less acid 4 per cent NaCl, e.g., pH 5.5 and above, the corrosion rate drops to a very low value, so small that its measurement presents difficulties by the weight-loss method.[17] For this situation the value of $\beta A/2.303bi$ is large, and dE_1/dpH therefore becomes small. It will be recognized that equation (5) essentially summarizes the conclusions of Gatty and Spooner and of D'Ans and Breckheimer, to the effect that for solutions in the more acid range of pH, dependence of potential on pH is large compared with a zero or lesser dependence at higher pH values.

The value of β for iron is equal to about 0.1,[16, 18] the average i from Stern's corrosion rates is equal to 1×10^{-5} amp/cm^2 in the pH range 1–4, and the corresponding dE_1/dpH in 4 per cent NaCl, as carefully determined by Stern, is -0.0577. Substituting these values in equation (5), b/A is found equal to 1.7×10^5 volts/amp/cm^2. The value of anodic overvoltage for iron in Fe^{++} solutions at low current densities is found to be of the order of 0.1 volt/0.1 ma/cm^2, based on apparent anodic area.[19] Therefore, the fraction A of total area representing the anode in the corrosion reaction is equal to 0.006. If b is larger, the value of A is larger, and vice versa; but it seems apparent, in any case, that the total anodic area at any moment in time is probably much less than the cathodic area, even though anodes and cathodes may interchange relative positions as the reaction proceeds. As mentioned previously, an increase of A by external anodic polarization produces a large percentage increase of available anodic surface, but the change is sluggish, accounting in part for the marked hysteresis of anodic polarization data.

The 10–13 per cent Cd amalgam electrode employed in the Weston Standard Cell is much less sensitive to change of hydrogen-ion activity than is iron, such effect as there is being accounted for by change in activity of the electrolyte by addition of acid.[20] The value of β for this electrode is again 0.1, i is approximately 2×10^{-9} amp/cm^2, and b is 0.013 volt/ma/cm^2.[9] Therefore, $dE_1/d\text{pH}$ is equal to $-3.5 \times 10^{-8}/A$. Since the apparent value of A for the amalgam is of the order of unity,[21] this absolute value of slope is several orders of magnitude smaller than the value for iron. The lower corrosion rate and the smaller anodic polarization account for the difference.

Change of Iron Potential with Ferrous-Ion Activity.—The change of E_a with ferrous-ion activity (Fe^{++}) for a hypothetical reversible iron electrode is equal to $0.0592/[2 \times 2.303\ (Fe^{++})]$. The change of corrosion potential, E_1, with (Fe^{++}) is calculated from equations (1) and (2) as follows:

$$E_a' = -E^\circ + \frac{0.0592}{2} \log (Fe^{++}) + \frac{bi}{A}, \tag{6}$$

where E° is the standard electrode potential and concentration polarization is again omitted.

$$\frac{dE_a'}{d(Fe^{++})} = \frac{0.0592}{2 \times 2.303\ (Fe^{++})} + \frac{b}{A} \frac{di}{d(Fe^{++})}, \tag{7}$$

and, from equation (1),

$$\frac{dE_c}{d(Fe^{++})} = -\frac{\beta}{2.303i} \frac{di}{d(Fe^{++})}. \tag{8}$$

Therefore, at the value of potential E_1 where i is the same for polarized anode and cathode,

$$\frac{dE_1}{d(Fe^{++})} = \frac{0.059}{2 \times 2.303\ (Fe^{++})} \frac{1}{1 + (2.303bi/\beta A)}. \tag{9}$$

From this equation, it is evident that $dE_1/d(Fe^{++})$ is small for large values of the local-action current i but approaches the theoretical value when i is small. In acid mediums, therefore, where i is large, the observed corrosion potential, E_1, does not respond to change of ferrous-ion activity contrary to the Nernst equation. However, in neutral or alkaline mediums, where the corrosion rate is much less, the correspondence is better. This again agrees with previously reported observations of the iron potential as a function of pH and ferrous-ion activity. The fact that values of the standard potential for iron calculated from potential measurements have in part succeeded can be attributed to the near-neutral electrolyte at the surface of the electrode after a long period of waiting which served to reduce the magnitude of local-action currents, thereby bringing the observed corrosion potential into closer correspondence with the reversible potential.

Using the previously listed numerical values, the term $1/(1 + 2.303bi/\beta A)$ in equation (9), representing the irreversible portion of $dE_1/d(Fe^{++})$, becomes equal to $1/(1 + 39)$, or 0.025, in the acid range of pH, and, therefore, the potential in this range is largely independent of ferrous-ion activity. In the neutral or alkaline range of

pH, applying a corrosion rate of 0.1 mdd, or $i = 4 \times 10^{-8}$ amp/cm², the same fraction becomes $1/(1 + 0.157) = 0.87$, indicating that $dE_1/d(Fe^{++})$ in this case approaches the reversible value.

For the 10–13 per cent Cd amalgam electrode, the irreversible factor is equal to $1/[1 + (6.0 \times 10^{-7}/A)]$, which, assuming $A = 1$, as discussed before, is practically indistinguishable from unity; hence this electrode is truly reversible. It is of interest to calculate from equation (6) that the difference in corrosion rates between electrodes used by Patrick and Thompson and by Randall and Frandsen necessary to account for the 0.031 volt difference in their reported $E°$ values is of the order of 0.5 mdd. This assumes that b/A was the same and did not vary with source of iron.

Change of Corrosion Rate with pH.—By setting polarized anode and cathode potentials equal to each other, an expression is obtained for the value of the local-action current i corresponding to the intersection of anode and cathode polarization curves. This value of i, as discussed previously, corresponds to the corrosion rate.

$$-0.0592 \text{ pH} - \beta \log \frac{i}{i_0} = -E_a + \frac{bi}{A}. \tag{10}$$

At very low values of i ($< 10^{-6}$ amp/cm²), the logarithm term predominates, and, except for the cathodic concentration-polarization term not included above, it is apparent that the logarithm of corrosion rate should be approximately a linear function of pH. But at higher values the corrosion rate itself is linear with pH. The smaller the value of exchange current i_0 and the greater the value of β, equivalent to high hydrogen overvoltage, the smaller is the corrosion rate. From equation (10), the change of i with pH is given by:

$$\frac{di}{d\text{pH}} = -\frac{0.0592}{(b/A) + (\beta/2.303i)},$$

or, in terms of the corrosion rate in mg/dm²/day,

$$\frac{d\text{mdd}}{d\text{pH}} = \frac{-1.5 \times 10^5}{(b/A) + [\beta/(9.2 \times 10^{-7} \text{ mdd})]}.$$

The numerical value of this expression in the pH range 1–4, where the corrosion rate approximates 25–30 mdd, is −0.86. This agrees satisfactorily with the data obtained by Stern and by Gemmell. In other words, the corrosion rate of iron in deaerated acids is changed very little by change of pH.

In near-neutral or alkaline mediums where the corrosion rate is of the order of 0.1 mdd, the change of corrosion rate with pH is still less. In these mediums, the corrosion rate presumably decreases to such a low value because the limiting cathodic current density for discharge of hydrogen ions, which decreases with decrease of H^+ activity or with increase of pH, attains an order of magnitude equal to that of the corrosion currents. This takes place in the region of pH above 5 or 6, causing a large increase in the concentration-polarization term $(RT/F) \ln [i_d/(i_d - i)]$, corresponding to hydrogen-ion activity at the electrode surface very much lower than that in the bulk of solution. The increased alkalinity is accompanied by precipitation of $Fe(OH)_2$ at the metal surface, which may act as a diffusion barrier

film, serving still further to maintain alkaline conditions at the surface and conforming with a reduced corrosion rate. The pH at the surface, for example, may achieve a value as high as 9.3, which is the pH of saturated $Fe(OH)_2$,[22] even though the bulk of solution is at a lower pH.

Discussion and Conclusions.—It is apparent that the above calculations for the irreversible behavior of iron must apply in principle to any corrodible metal. Deviation from reversibility is characteristic in particular of the transition metals, which, as a group, are noted for their high values of anodic polarization or anodic overvoltage. For nontransition metals, e.g., zinc and copper, reversibility is better because anodic polarization is less, and, in the case of copper, the corrosion rate is small compared with that of iron. The satisfactory thermodynamic behavior of amalgams, e.g., Zn and Cd, is accounted for by both low anodic polarization and small corrosion rates, the low rates of attack in deaerated aqueous mediums resulting from the high hydrogen overvoltage of mercury. The presence of dissolved oxygen in electrolytes depolarizes the cathodes and hence increases local-action currents, which, in turn, shift the observed potential in the cathodic direction. Reversibility is better, therefore, in electrolytes free of dissolved oxygen or other depolarizers.

It is concluded:

1. The iron potential is irreversible. The measured value depends on the magnitude of local-action currents accompanying corrosion by the electrolyte and on the extent of anodic polarization. This factor accounts for the observed dependence of the iron potential on pH. Local-action currents vary with all the causes of varying corrosion rate, including trace impurities in the metal, dissolved oxygen in the electrolyte, and composition and pH of the electrolyte.

2. The observed potential $E_{obs.}$ is always more cathodic than the reversible or thermodynamic value $E_{therm.}$, according to the relation

$$E_{therm.} = E_{obs.} + f(i),$$

where $f(i)$ is a function of local-action currents calculable from polarization data and reaction rates.

3. Transition-metal potentials, especially, deviate from reversibility because of characteristically high values of anodic polarization compared with nontransition metals, and by reason of their appreciable reaction rate with aqueous solutions of their salts.

4. It is feasible to calculate only approximate standard potentials for iron and for similar metals from assumed equilibrium potentials.

[1] *J. Am. Chem. Soc.*, **46**, 89 (1924).

[2] O. Gatty and E. Spooner, *The Electrode Potential Behavior of Corroding Metals in Aqueous Solutions* (Oxford University Press, 1938), p. 312.

[3] J. D'Ans and W. Breckheimer, *Zs. f. Elektrochem.*, **56**, 585 (1952).

[4] K. Bonhoeffer and W. Jena, *Zs. f. Elektrochem.*, **55**, 151 (1951).

[5] W. H. Hampton, *J. Phys. Chem.*, **30**, 980 (1926).

[6] M. Randall and M. Frandsen, *J. Am. Chem. Soc.*, **54**, 47 (1932).

[7] W. Patrick and W. Thompson, *J. Am. Chem. Soc.*, **75**, 1184 (1953).

[8] R. B. Mears and R. Brown, *J. Electrochem. Soc.*, **97**, 75 (1950).

[9] H. H. Uhlig, *J. Electrochem. Soc.*, **100**, 173 (1953).

[10] S. Glasstone, *Electrochemistry* (New York: D. Van Nostrand Co., 1947), p. 448.

[11] J. A. V. Butler, *Electrical Phenomena at Interfaces* (London: Methuen & Co., 1951), p. 198.

[12] G. Gemmell, M.I.T. thesis, 1953.

[13] C. Wagner and W. Traud, *Zs. f. Elektrochem.*, **44**, 391 (1938).

[14] Butler, *op. cit.*, p. 201.

[15] Glasstone, *op. cit.*, p. 457.

[16] M. Stern, unpublished data, Corrosion Laboratory, M.I.T.

[17] G. Gemmell found that the corrosion rate for mild steel (0.06 per cent C) in deaerated 4 per cent NaCl of pH 5.6 was zero within the experimental variations of such measurements, whereas at pH 3.8 the rate averaged 24 mg/dm²/day (mdd), in agreement with Stern's value for iron. This means that the rate at pH 5.6 must be less than perhaps the order of 0.5 mdd. Pryor and Cohen, *J. Electrochem. Soc.*, **100**, 206 (1953), report a value of 0.3 mdd for the corrosion rate of mild steel in deaerated water.

[18] Gatty and Spooner, *op. cit.*, p. 318.

[19] W. Latimer and J. Hildebrand, *Reference Book of Inorganic Chemistry* (New York: Macmillan Co., 1941), p. 473.

[20] C. Wagner, *J. Electrochem. Soc.*, **100**, 524 (1953).

[21] Based on atomic percentages and atomic radii of Cd and Hg, the fraction of surface composed of Cd atoms for 10–13 per cent Cd amalgam is equal to about 0.2. However, in measurements of anodic polarization, this fraction remains constant up to fairly high current densities because the potential for formation of Hg_2^{++} is not approached so long as the surface concentration of Cd is appreciable. Therefore, measured values of b are for an anodic area similar to that which applies to the corrosion process, and, hence, the apparent value of A in eq. (2) under these conditions is equal to unity. Data by Wagner and Traud, *op. cit.*, for zinc amalgam confirm this conclusion. For iron, on the other hand, the anodic surface becomes larger, the higher the value of current density externally applied for anodic polarization, and, hence, A is usually less than unity.

[22] D. Leussing and I. Kolthoff, *J. Am. Chem. Soc.*, **75**, 2476 (1953).

---

*THE NATURAL C^{14} CONTENTS OF MATERIALS FROM HARD-WATER LAKES**

BY EDWARD S. DEEVEY, JR., MARSHA S. GROSS, G. E. HUTCHINSON, AND HENRY L. KRAYBILL

YALE UNIVERSITY, NEW HAVEN, CONN.

Communicated March 17, 1954

In the ordinary procedure of radiocarbon dating, as developed by Libby and his co-workers, it is assumed that the carbon, before entering the material under investigation, has achieved isotopic exchange equilibrium with the CO_2 of the air. It has been pointed out, notably by Godwin,[1] that this assumption may well be false in the case of material formed in hard waters, either by photosynthetic fixation of carbon in organic matter or by precipitation of $CaCO_3$. Ordinary hard waters contain a considerable quantity of calcium bicarbonate, formed by the action on limestone of rain or ground water containing free CO_2. The free CO_2 initially present in such waters may be regarded as being in isotopic exchange equilibrium with the CO_2 of the atmosphere; the limestone, however, is usually very old compared with the half-life of C^{14}. The initial bicarbonate formed may therefore be expected to have about half the specific activity of modern wood or other substances formed from atmospheric CO_2. Exchange according to the scheme

$$HCO_3^- \rightleftharpoons H_2CO_3 \rightleftharpoons CO_2\ (aq.) \rightleftharpoons CO_2\ (gas)$$

may then be expected, and an approximately uniform distribution of C^{14} will ulti-

mately be reached. Since, however, at equilibrium an enrichment of C^{13} of the order of 3 per cent occurs in the carbon of the HCO_3^- ion compared with the CO_2 in the gaseous phase,[2] an enrichment of C^{14}, probably of the order of 6 per cent, is to be expected.[3]

It is known that some of the steps in the process outlined above are not instantaneous. Moreover, Ohle[4] has shown that in calcareous waters there is apt to be a considerable amount of colloidal $CaCO_3$ present. This colloidal carbonate is extremely stable in the absence of excess CO_2, even in the presence of the solid phase. It is probably formed as the result of the photosynthetic withdrawal of CO_2 from bicarbonate solutions. If this process occurred in the drainage basin, or at the surface of mud containing appreciable amounts of old carbonate, the colloidal $CaCO_3$ could provide a reservoir of carbon of low activity even in a water in which complete isotopic equilibrium existed between the components in ionic and true solution and in the gaseous phase.

In view of these theoretical possibilities and of the fact that Blau, Deevey, and Gross[5] obtained evidence of spurious antiquity for a marl deposited in Pyramid Valley, North Canterbury, New Zealand, in what must once have been a shallow hard-water lake, it has seemed desirable to examine the problem in greater detail.

Most of our samples are derived from Queechy Lake, New York, a small lake of area 60.8 ha., of maximum depth 13.6 meters, and of mean depth 6.9 meters, lying on the Stockbridge marble, a metamorphosed Paleozoic limestone. The lake lacks inlets and receives its water by seepage through the limestone and the overlying glacial drift, which itself contains much of the country rock. A typical "marl girdle" is present in the shallow parts of the basin, and a considerable development of a benthic blue-green algal mat has been noted. We have also studied two samples collected by R. J. Benoit in Lake Zoar, Connecticut, an artificial lake outside the region of outcrop of the limestone, but receiving its water from the Housatonic River, which drains the limestone north and west of New Milford, Connecticut. The bicarbonate alkalinity of both these lakes is given by Deevey[6] as 123 mg. HCO_3 per liter.

Several samples of the vegetation of soft-water lakes and bogs in Connecticut have been used as controls, in addition to the modern wood samples employed as standards. The results given in Table 1 are in counts per minute, corrected for carbon content. They refer to a layer of pure carbon of effectively infinite thickness and of area 410 $cm.^2$, in practice weighing about 10 gm. They are not otherwise corrected for the instrumental characteristics of our counters, but comparison with the work of other investigators may be made by means of our value for modern wood. The error appended to each determination is the standard error due to counting statistics only. The error appended to each average value refers to the standard deviation of that mean value, as determined by the scatter of the individual results. The dissolved, colloidal, and suspended carbonate of the water of Lake Queechy was obtained by acidifying 650 liters of the lake water and collecting the evolved gas in ascarite. The carbon of the bicarbonate and carbonate in this water has the specific activity of a solution, formed by the action of atmospheric CO_2 on old limestone, that has proceeded halfway to exchange equilibrium. The carbon of the modern lake water therefore has a spurious "age" of 2,200 years. All the other carbon-containing materials actually produced in the lake have low

specific activities, but none quite so low as the water. It is possible that in the sample studied there was some CO_2 from colloidal or suspended $CaCO_3$ with a rather lower specific activity than that of the bicarbonate-ion carbon present. The products of photosynthetic fixation and deposition of $CaCO_3$ are almost as low, however, which indicates that the carbon of the bicarbonate ion present in solution must have a considerably lower specific activity than that of the carbon of atmospheric CO_2.

TABLE 1

Sample	Description	C^{14} (c.p.m.)
Blanks	Anthracite	-0.00 ± 0.08
	Anthracite	$-.09 \pm .08$
	Paleozoic limestone	$-.10 \pm .09$
	Paleozoic limestone	-0.06 ± 0.11
Modern wood:		
Y-239	Black birch (*Betula lenta*), 10–20 years old, Ansonia, Conn.	$+5.96 \pm 0.11$
Y-171	Black oak (*Quercus velutina*), Ansonia, Conn.	$+5.78 \pm .11$*
Y-171	Black oak (*Quercus velutina*), Ansonia, Conn.	$+6.08 \pm .11$*
Y-172	Black oak (*Quercus velutina*), Ansonia, Conn.	$+5.80 \pm .11$*
Y-172	Black oak (*Quercus velutina*), Ansonia, Conn.	$+5.70 \pm 0.10$*
Average of modern woods		$+5.86 \pm 0.08$
Soft-water localities in Connecticut, emergent vegetation:		
Y-29	*Chamaedaphne calyculata*, Old Man MacMullen's Pond, Canaan	$+6.33 \pm 0.10$
Y-29	*Chamaedaphne calyculata*, Old Man MacMullen's Pond, Canaan	$+5.99 \pm .12$
Y-30	*Vaccinium* sp., Old Man MacMullen's Pond, Canaan	$+6.20 \pm .08$
Soft-water localities in Connecticut, submerged vegetation:		
Y-63	*Sphagnum* spp., Bethany Bog	$+6.03 \pm .10$
Y-56	Mixed submerged phanerogams, Lake Quassapaug	$+6.44 \pm .10$
Y-56	Mixed submerged phanerogams, Lake Quassapaug	$+5.93 \pm 0.12$
Average of soft-water plants		$+6.15 \pm 0.08$
Queechy Lake:		
Y-145-6	CO_2 from bicarbonate and carbonate in water	$+4.60 \pm 0.10$†
Y-145-3	*Chara* sp., organic carbon	$+5.07 \pm .07$
Y-145-3	*Chara* sp., inorganic calcareous material	$+5.02 \pm .10$
Y-145-4	*Potamogeton* sp., organic carbon	$+4.99 \pm .11$
Y-145-4	*Potamogeton* sp., organic carbon	$+5.18 \pm .16$
Y-145-5	*Potamogeton* sp., inorganic calcareous incrustation	$+5.02 \pm .11$
Y-145-5	*Potamogeton* sp., inorganic calcareous incrustation	$+5.09 \pm .11$
Y-145-2	*Elliptio* (*E.*) *complanatus*, soft parts	$+4.83 \pm .11$
Y-145-2	*Elliptio* (*E.*) *complanatus*, soft parts	$+4.82 \pm .08$
Y-145-1	*Elliptio* (*E.*) *complanatus*, shell	$+5.20 \pm .14$
Y-145-1	*Elliptio* (*E.*) *complanatus*, shell	$+4.87 \pm .07$
Y-224	Deep-water mud	$+4.68 \pm .09$
Y-224	Deep-water mud	$+4.97 \pm .15$
Y-21	*Scirpus* sp., emergent	$+6.00 \pm .10$
Y-22	*Carex* cf. *lasiocarpa*, emergent, from shallow water	$+6.11 \pm 0.08$
Lake Zoar:		
Y-20	*Hydrodictyon reticulatum*	$+5.44 \pm 0.11$
Y-27	*Elodea* sp.	$+5.35 \pm 0.09$

* Previously reported in Blau, Deevey, and Gross, *Science*, **118**, 1–6 (1952).
† Determined by Dr. Hans E. Suess and expressed as 77 ±2 per cent of the value for modern wood. The mean value for modern wood is taken here as the mean of all modern woods and soft-water samples, namely, 6.02.

There is clearly no difference between the activity of the organic carbon in the

submerged water plants of the lake and that of the carbon of the carbonate deposit associated with these plants. Both *Chara* and *Potamogeton* doubtless use bicarbonate ions as a carbon source, and the fact that both the organic and carbonate carbon are derived from the same material presumably underlies the identity in C^{14} content.

The two emergent plants growing rooted in the shallow waters of the lake, namely, *Scirpus* and *Carex*, clearly obtained their carbon directly from the atmosphere.

The Lake Zoar samples, though somewhat richer in C^{14} than those from Queechy Lake, are still deficient in C^{14} when compared with modern wood. Evidently, isotopic exchange equilibrium is not fully achieved in a relatively shallow river over a course of 15 miles.

The first runs on the samples Y-56 and Y-29 (Table 1) appear to be abnormally high compared to the two later runs. This could be caused by slight contamination; however, no independent evidence of contamination was recorded at the time. The mean for all the soft-water plants is 6.15 c.p.m. including, or 6.04 c.p.m. excluding, these samples. Such values are comparable to that from the modern wood of black birch studied during the present investigation but are slightly higher than some of our earlier wood samples. The significance of this slight difference is under investigation.

It is apparent that if marl, shell, or water plants that developed in an ancient lake geochemically comparable to Queechy were used in radiocarbon dating, spurious ages, up to 2,000 years in excess of the true values, would be obtained. The rejection of the Pyramid Valley dates based on lake marl, as suggested by Blau, Deevey, and Gross, is obviously justified. Great care must clearly be exercised in the future when material from highly calcareous regions is being examined. It seems probable that in alkaline lakes in closed basins in semiarid regions not underlain by limestones, all the carbon in the carbonate and bicarbonate in the water is of atmospheric origin. Samples from such lakes are likely to be free from the error here discussed.

We are greatly indebted to Dr. Hans E. Suess for determining the activity of the dissolved bicarbonate from Queechy Lake by acetylene gas counting, our supply of material being inadequate for use as carbon in our solid sample counter. Percy A. Morris, of the Peabody Museum, Yale University, kindly identified the clams from Queechy Lake as *Elliptio* (*E.*) *complanatus*. We also wish to express our appreciation of the work of Duane Malm and of Kenneth Buckley, bursary assistants in the Geochronometric Laboratory.

* Contribution from the Geochronometric Laboratory and from the Osborn Zoölogical Laboratory of Yale University. Supported by grants from the Rockefeller Foundation and the Office of Naval Research.

[1] H. Godwin, *Am. J. Sci.*, **249**, 301–307 (1951).

[2] A. O. Nier and E. A. Gulbranson, *J. Am. Chem. Soc.*, **61**, 697–698 (1939); B. F. Murphey and A. O. Nier, *Phys. Rev.*, **59**, 771–772 (1941); H. Craig, *Geochim. cosmochim. acta*, **3**, 53–92 (1953).

[3] E. C. Anderson and W. F. Libby, *Phys. Rev.*, **81**, 64–69 (1951).

[4] W. Ohle, *Arch. f. Hydrobiol.*, **46**, 153–285 (1952).

[5] M. Blau, E. S. Deevey, Jr., and M. S. Gross, *Science*, **118**, 1–6 (1952).

[6] E. S. Deevey, Jr., *Am. J. Sci.*, **238**, 717–741 (1940).

BOUNDEDNESS AND STATIONARITY OF TIME SERIES

By S. Bochner

DEPARTMENT OF MATHEMATICS, PRINCETON UNIVERSITY

Communicated March 17, 1954

The aim of this paper is to demonstrate Theorem 4[1] and also Theorem 5.

We take in $-\infty < t < \infty$ an operator

$$\Lambda x(t) = \sum_{\rho=1}^{r} c_\rho x(t + \omega_\rho) \tag{1}$$

with constant coefficients c_ρ and arbitrary real spans ω_ρ, and we are introducing the function

$$T(\alpha) = \sum_{\rho=1}^{r} c_\rho e^{i\omega_\rho \alpha}, \tag{2}$$

which is such that

$$\Lambda_t e^{i\alpha t} = T(\alpha) e^{i\alpha t}, \quad -\infty < \alpha < \infty. \tag{3}$$

The real zeros of $T(\alpha)$, if any, will be denoted by $\{\alpha_n\}$, always.

We now take a fixed Hilbert space H and functions $x(t)$ in $(-\infty, \infty)$ which have values in H and are measurable. Our main concern will be bounded functions,

$$\|x(t)\| < M, \tag{4}$$

but it is appropriate to introduce the more general class of functions for which

$$\int_{-\infty}^{\infty} \frac{\|x(t)\|^2}{1 + t^2}\, dt < \infty. \tag{5}$$

We denote the class by P_1, and we will say that $x_m(t)$ is P_1-convergent to $x(t)$ if

$$\lim_{m \to \infty} \int_{-\infty}^{\infty} \frac{\|x_m(t) - x(t)\|^2}{1 + t^2}\, dt = 0.$$

If $x(t) \in P_1$, then the Fejer sequence

$$x_m(t) = \frac{2}{\pi m} \int_{-\infty}^{\infty} \frac{[\sin (m/2)\,(t - \tau)]^2}{(t - \tau)^2}\, x(\tau)\, d\tau \tag{6}$$

is P_1-convergent to it. If $x(t)$ is bounded and uniformly continuous, then this sequence is even uniformly convergent.

A function $x(t)$ in P_1 has a (generalized) Fourier integral representation

$$x(t) \sim \int_{-\infty}^{\infty} e^{i\alpha t}\, dE(\alpha), \tag{7}$$

in which $E(\alpha)$ has values in H and is determined up to a constant and may be defined by

$$2\pi i E(\alpha) = \int_{-1}^{1} x(t)\, \frac{e^{-i\alpha t} - 1}{-it}\, dt + \left(\int_{-\infty}^{-1} + \int_{1}^{\infty} \right) x(t)\, \frac{e^{-i\alpha t}}{-it}\, dt. \tag{8}$$

The second integral need not be absolutely convergent, but it is a Plancherel limit-in-mean, and the function $E(\alpha)$ is defined almost everywhere and is square-integrable over any finite interval at any rate. Also, if $\varphi(t)$ is a numerical function with

$$\int_{-\infty}^{\infty} \frac{|\varphi(t)|^2}{1+t^2}\,dt < \infty$$

and if we put

$$\psi(\alpha) = \int_{-\infty}^{\infty} e^{i\alpha\tau}\,\varphi(\tau)\,d\tau, \tag{9}$$

then we have

$$\int_{-\infty}^{\infty} \varphi(\tau)x(\tau)\,d\tau = \int_{-\infty}^{\infty} \psi(\alpha)dE(\alpha) \tag{10}$$

and, more generally,

$$\int_{-\infty}^{\infty} \varphi(\tau - t)x(\tau)\,d\tau = \int_{-\infty}^{\infty} e^{i\alpha t}\,\psi(\alpha)dE(\alpha). \tag{11}$$

The symbolic integral on the right-hand side may be evaluated by

$$\int_a^b \psi(\alpha)dE(\alpha) = \psi(b)E(b) - \psi(a)E(a) - \int_a^b \psi'(\alpha)E(\alpha)d\alpha, \tag{12}$$

so that, for instance, for the functions (6), we have

$$x_m(t) = \int_{-m}^{m} \left(1 - \frac{|\alpha|}{m}\right) e^{i\alpha t}\,dE(\alpha) \tag{13}$$

in this sense.

The symbolic entity "$dE(\alpha)$" will be called the "spectrum" of $x(t)$. It is zero in an open interval (α', α''), $dE = 0$, if $E(\alpha) =$ constant there, and it is zero at a point β if it is zero in an interval $(\beta - \epsilon, \beta + \epsilon)$. It has an isolated point β, if $E(\alpha) = b'$ in $(\beta - \epsilon, \beta)$ and b'' in $(\beta, \beta + \epsilon)$, and the saltus $b = b'' - b'$ is then such that $be^{i\beta t}$ is the part of integral (11) in $(\beta - \epsilon, \beta + \epsilon)$. If the spectrum has only isolated points, $\{\beta_n\}$, we can put symbolically

$$x(t) \sim \sum b_n e^{i\beta_n t}, \tag{14}$$

and in this case (13) is the exponential polynomials

$$x_m(t) = \sum_{|\beta_n| < m} \left(1 - \frac{|\beta_n|}{m}\right) b_n e^{i\beta_n t}. \tag{15}$$

Thus $x(t)$ is then the P_1-limit of the trigonometric sums (15), but $x(t)$ is not necessarily almost periodic in a conventional sense. However, if $x(t)$ is also known to be bounded and uniformly continuous, then it is almost periodic in the strict sense of H. Bohr (for functions with values in H, that is), and (14) is then its Fourier expansion.

If we now consider the relation

$$\Lambda x(t) = y(t) \tag{16}$$

for the function (7), then $y(t)$ is likewise in P_1,

$$y(t) \sim \int_{-\infty}^{\infty} e^{i\alpha t}\,dF(\alpha), \tag{17}$$

and (3) implies the relations

$$T(\alpha)dE(\alpha) = dF(\alpha), \tag{18}$$

in the sense that for any differentiable function $\psi(\alpha)$ we have

$$\int_{\alpha'}^{\alpha''} \psi(\alpha)T(\alpha)dE(\alpha) = \int_{\alpha'}^{\alpha''} x|\,\alpha|\,dF(\alpha), \tag{19}$$

in accordance with (12). This implies that if we have $T(\alpha) \neq 0$ in an open interval (α', α''), then we also have there

$$dE(\alpha) = \frac{1}{T(\alpha)}\,dF(\alpha), \tag{20}$$

so that $dF(\alpha) = 0$ implies $dE(\alpha) = 0$ there, and hence the following conclusion.

THEOREM 1. *If $x(t) \in P_1$ satisfies the homogeneous equation*

$$\Lambda x(t) = 0, \tag{21}$$

then we have

$$x(t) \sim \sum a_n e^{i\alpha_n t}, \tag{22}$$

and if $x(t)$ is bounded and uniformly continuous, it is also almost periodic.

Less obvious is the following statement:

THEOREM 2. *If $x(t) \in P_1$ is a solution of the nonhomogeneous equation (16) and if the spectrum $dF(\alpha)$ of $y(t)$ is isolated at a point α_n, then, of necessity, $dF(\alpha) = 0$ at α_n.*

Thus, if the spectrum of $y(t)$ reduces to the isolated points $\{\alpha_n\}$ at most, then $y(t) = 0$ and $x(t)$ is as in Theorem 1.

Proof: If we apply (19) to $\alpha' = \alpha_n - \epsilon$, $\alpha'' = \alpha_n + \epsilon$, and define $\psi(\alpha)$ to be equal to 1 in (α', α'') and to be equal to 0 in $(-\infty, \alpha' - \epsilon)$ and $(\alpha'' + \epsilon, \infty)$, then we can introduce two new functions,

$$x^\circ(t) \sim \int e^{i\alpha t}dE^\circ(\alpha),\ y^\circ(t) \sim \int e^{i\alpha t}dF^\circ(\alpha), \tag{23}$$

which are such that

$$dE^\circ(\alpha) = \psi(\alpha)dE(\alpha), \qquad dE^\circ(\alpha) = \psi(\alpha)dF(\alpha) \tag{24}$$

in $-\infty < \alpha < \infty$. We again have

$$\Lambda x^\circ(t) = y^\circ(t), \tag{25}$$

and $dF^\circ(\alpha) = 0$ except at the isolated point α_n, at which its value b is the same as it was for $y(t)$. Thus we have $y^\circ(t) = be^{i\alpha_n t}$. Now (18) implies $x^\circ(t) = ae^{i\alpha_n t}$, and from (3) we now conclude that $b = 0$, as claimed.

Definition 1: We will say that $y(t)$ is K-stationary (K for "Khintchine") if the covariance function

$$\gamma(r, s) = [y(r), y(x)] \tag{26}$$

is a function of $t = r - s$ only, $\gamma(r, s) = \gamma(r - s)$.

Such a $y(t)$ is constant in norm and uniformly continuous, and thus, as a function of P_1, it has an expansion (17). Now it is known that, in this case, $F(\alpha)$ has right and left limits $F(\alpha \pm 0)$, and if we normalize it to be continuous from the left, say, and if with each bounded interval $\Delta: \alpha' \leq \alpha < \alpha''$ we associate the value $F(\Delta) =$

$F(\alpha'') - F(\alpha')$, then for any two nonoverlapping intervals Δ_1, Δ_2 we have the orthogonality property

$$[F(\Delta_1), \quad F(\Delta_2)] = 0. \tag{27}$$

Therefore, the interval function

$$\Gamma(\Delta) = \|F(\Delta)\|^2 \geq 0 \tag{28}$$

is additive, and, in fact, we have

$$\gamma(t) = \int_{-\infty}^{\infty} e^{i\alpha t}\, d\Gamma(\alpha), \tag{29}$$

$$\int_{-\infty}^{\infty} d\Gamma(\alpha) = \gamma(0) < \infty. \tag{30}$$

Next, if $x(t)$ is bounded, $\|x(t)\| \leq N$, then, for any numerical $\varphi(t)$ with

$$\int_{-\infty}^{\infty} |\varphi(t)|\, dt < \infty,$$

we can form

$$\xi_\varphi(t) = \int_{-\infty}^{\infty} \varphi(t)x(t)\, dt, \tag{31}$$

and we have

$$\|\xi_\varphi\| \leq N \int_{-\infty}^{\infty} |\varphi(t)|\, dt.$$

But we will stipulate a sharper property of boundedness, as follows:

Definition 2: We say that $x(t)$ is V-bounded (V for "variation") if $x(t)$ is bounded, and we have

$$\left\| \int_{-\infty}^{\infty} \varphi(t)\, x(t)\, dt \right\| \leq M \cdot \sup_\alpha |\psi(\alpha)|, \tag{32}$$

where $\psi(\alpha)$ is the transform (9).

THEOREM 3. *Any K-stationary function is V-bounded and so is more generally any (nonstationary) function $x(t)$ whose covariance function $\gamma(r, s)$ has a representation*

$$\gamma(r, s) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{i(\alpha r - \beta s)}\, d\Gamma(\alpha, \beta),$$

with

$$\int \int |d\Gamma(\alpha, \beta)| = M^2 < \infty.$$

(M. Loève's harmonizable functions.)

In fact, we have

$$\|\xi_\varphi\|^2 = \left[\int \varphi(r)\, x(r)\, dr, \quad \int \varphi(s)\, x(s)\, ds\right] =$$
$$\int \int \varphi(\alpha)\overline{\varphi(s)}\gamma(r, s)\, dr\, ds = \int \int \int \int \varphi(r)\overline{\varphi(s)} e^{i(\alpha r - \beta s)}\, dr\, ds\, d\Gamma(\alpha, \beta) =$$
$$\int \int \psi(\alpha)\overline{\psi(\beta)}\, d\Gamma(\alpha, \beta) \leq M^2 \sup_\alpha |\psi(\alpha)|^2,$$

as claimed.

Definition 3: We denote by R the open set on the line $(-\infty, \infty)$ which arises by deletion of the zeros $\{\alpha_n\}$ of $T(\alpha)$.

THEOREM 4. *If, in relation (16), the primary function $x(t)$ is V-bounded and the*

smoothed function $y(t)$ is even K-stationary, then, for the spectrum of $y(t)$, we have

$$\int_R \frac{d\Gamma(\alpha)}{|T(\alpha)|^2} = M(< \infty), \tag{33}$$

and its saltuses $b_n = F(\alpha_n + 0) - F(\alpha_n - 0)$ are all zero.

There is a K-stationary function $x^1(t)$ with the spectrum

$$x^1(t) \sim \int_R \frac{e^{i\alpha t}}{T(\alpha)} dF(\alpha), \tag{34}$$

and it is again a solution of $\Lambda\, x^1(t) = y(t)$, and it differs from the primary function $x(t)$ by a difference $x^2(t) = x(t) - x^1(t)$ of the special form (22), which, if $x(t)$ is uniformly continuous, is even almost periodic.

Proof: We again go over from relation (16) to relation (25) by means of a "multiplier" $\psi(\alpha)$, but now the latter is of the following general form: It is defined and differentiable in $(-\infty, \infty)$; it vanishes outside some finite interval and also in a neighborhood of each point α_n; and we have $|\psi(\alpha)| \leq 1$.

The function $y^\circ(t)$ is again K-stationary, and we have

$$T(\alpha) dE^\circ(\alpha) = dF^\circ(\alpha),$$

but at present we can hence obtain

$$dE^\circ(\alpha) = \frac{dF^\circ(\alpha)}{T(\alpha)}$$

for α everywhere in $(-\infty, \infty)$, and thus $x^\circ(t)$ is likewise K-stationary. Therefore, (32) implies

$$\int_{-\infty}^{\infty} \frac{d\Gamma^\circ(\alpha)}{|T(\alpha)|^2} \equiv \int_R |\psi(\alpha)| \frac{d\Gamma(\alpha)}{|T(\alpha)|^2} \leq M,$$

and, if we vary $\psi(\alpha)$ according to its possibilities, we hence obtain (33). Therefore, there exists a K-stationary function (34), and we have for it

$$\Lambda x^1(t) \sim \int_R e^{i\alpha t} dF(\alpha).$$

Thus, for the difference $x^2(t)$, we obtain

$$\Lambda x^2(t) \sim \sum b_n e^{i\alpha_n t},$$

and by Theorem 2 the right side must be zero. Hence the theorem.

We now take an interval function $X(\Delta)$, with values in H, which need not be the indefinite integral of a point function $x(t)$, and we are stating the following definitions which were introduced in another context.[2]

Definition 4: We say that $X(\Delta)$ is $L_{2,2}$-finite if there is an M such that we have

$$\left\| \int_{-\infty}^{\infty} \varphi(t) dX(t) \right\|^2 \leq M^2 \int_{-\infty}^{\infty} |\varphi(t)|^2\, dt,$$

at first for step functions $\varphi(t)$ and then, by extension, for all φ in L_2. We call $X(\Delta)$ a Khintchine function, if, for $\varphi_1, \varphi_2 \in L_2$, the quantity

$$\left[\int \varphi_1(r)\, dX(r), \quad \int \varphi_2(s)\, dX(s)\right] \tag{35}$$

is the same if we replace $\varphi_1(r)$, $\varphi_2(s)$ by $\varphi_1(r+t)$, $\varphi_2(s+t)$ for any given real t. And we call $X(\Delta)$ orthogonal if (38) is 0 for any two continuous functions $\varphi_1(r)$, $\varphi_2(s)$ for which

$$\int_{-\infty}^{\infty} \varphi_1(t)\ \overline{\varphi_2(t)}\ dt = 0.$$

Now we have shown[2] that an $L_{2,2}$-finite $X(\Delta)$ has a Fourier transform

$$\frac{dX(t)}{dt} \sim \int_{-\infty}^{\infty} e^{i\alpha t} dE(\alpha),$$

in which $E(\Delta)$ is again $L_{2,2}$-finite, so that, as before,

$$\int_{-\infty}^{\infty} \varphi(t) dX(t) \sim \int_{-\infty}^{\infty} \psi(\alpha) dE(\alpha).$$

Furthermore, $X(\Delta)$ is a Khintchine function if and only if $E(\Delta)$ is orthogonal; and if $E(\Delta)$ is orthogonal, then its covariance $\Gamma(\Delta)$, although not finite over the total line $(-\infty, \infty)$, is absolutely continuous and is, in fact, the indefinite integral of a point function which is bounded in $(-\infty, \infty)$. The effect of the latter property is that $E(\Delta)$ does not have isolated points, and the following very precise statement can be proved.

THEOREM 5. *If a primary function $X(\Delta)$ is $L_{2,2}$-finite and the smoothed function*

$$\sum_{\rho=1}^{r} c_\rho\, X(\Delta + \omega_\rho) = Y(\Delta) \tag{36}$$

is even a Khintchine function, then $X(\Delta)$ is likewise a Khintchine function.

A Wiener process is an $L_{2,2}$-finite function $X(\Delta)$ which is symmetric-Gaussian, Khintchine, and orthogonal, all at the same time. In the smoothing operation (36) the Khintchine and symmetric-Gaussian character is preserved both ways; the orthogonality, however, is not; and orthogonality means stochastic independence of the random variables $X(\Delta)$ for nonoverlapping intervals in a weak sense.

[1] For the general method of this paper see my papers "Ueber gewisse Differential- und allgemeinere Gleichungen deren Lösungen fast periodisch sind I, II, III," *Math. Ann.*, **102,** 409–504 (1929); **103,** 588–597 (1930); **104,** 579–589 (1931); and also my book *Fouriersche Integrale* (Chelsea, 1948).

[2] In a forthcoming book entitled *Harmonic Analysis and the Theory of Probability.* See also my paper "Fourier Transforms of Time Series," these PROCEEDINGS, **39,** 302–307 (1953).

ANALYTIC ALMOST-PERIODIC FUNCTIONS. I

BY OSCAR GOLDMAN

BRANDEIS UNIVERSITY

Communicated by Oscar Zariski, March 5, 1954

Let G be a compact, connected Abelian group, h a continuous homomorphism of the group R of positive real numbers into G, such that $h(R)$ is dense in G. We shall define a notion of analytic function in $M = R \times G$.

If $z = x + iy$ is a complex number, we set $e(z) = [\exp x, h(\exp y)]$. It is readily

verified that e is a continuous homomorphism of the additive group of complex numbers into M with a dense image.

If U is an open set in M, f a continuous complex-valued function defined on U, we shall say that f is *analytic* on U, if, for every $\alpha \in M$, the function $f[\alpha e(z)]$ is analytic in the open set $e^{-1}(\alpha^{-1}U)$.

Analytic functions on M may be obtained by means of the characters of G. If χ is a character of G, $r \to \chi h(r)$ is a character of R, so that there exists a real number $c(\chi)$ with $\chi h(r) = r^{ic(\chi)}$. Then, as may readily be verified, the function X defined on M by $X(r, u) = r^{c(\chi)}\chi(u)$ is analytic. The set of these functions is sufficiently large to separate the points of M.

If f is defined on an open set, U, of M, we may define (when they exist) two derivatives of f as follows:

$$f_1'(r, u) = \frac{d}{dx} f(re^x, u)|_{x=0}, f_2'(r, u) = \frac{d}{dy} f[r, uh(e^y)]|_{y=0}.$$

THEOREM 1. *Let f be a function analytic on an open set U. Then f_1' and f_2' both exist in U, are continuous there, and $f_1' + if_2' = 0$. Conversely, if f is a continuous function on U for which f_1' and f_2' exist and are continuous in U, and $f_1' + if_2' = 0$, then f is analytic on U.*

We prove first the converse. For $\alpha \in M$, set $U_\alpha = e^{-1}(\alpha^{-1}U)$, and $F_\alpha(z) = f[\alpha e(z)]$, $z \in U_\alpha$. Then, $(\partial/\partial x)F_\alpha(z) = f_1'[\alpha e(z)]$, $(\partial/\partial y)F_\alpha(z) = f_2'[\alpha e(z)]$, which shows that $F_\alpha(z)$ satisfies the Cauchy-Riemann equation in U_α, so that $F_\alpha(z)$, and hence f is analytic.

Assuming that f is analytic in U, $F_\alpha(z)$ is analytic in U_α; the previous computation shows that f_1' and f_2' both exist and that $f_1' + if_2' = 0$ in U. We need then only prove the continuity of these derivatives. To do this, let $\alpha_0 \in U$, and let U_0 be an open set containing α_0, whose closure is compact and contained in U. Now there exists an open, connected set D in the complex plane, containing 0, such that $\alpha_0 e(D)$ is contained in U_0. Let C be a simple closed curve contained in D, enclosing 0, and let D' be a compact, connected subset of D containing in its interior C as well as the region enclosed by C. Then $\alpha_0 e(D')$ is a compact subset of U_0. Since f is uniformly continuous on U_0, given any positive ϵ, there exists a neighborhood N of 1 in M, such that $|f(\alpha) - f(\beta)| < \epsilon$, whenever α and β are in U_0 and $\alpha\beta^{-1}$ is in N.

Because $\alpha_0 e(D')$ is a compact subset of the open set U_0, we can find an open set N', containing 1 and contained in N, such that $\alpha_0 \gamma N' \subset U_0$, for every γ in $e(D')$. Now let β be an element of $\alpha_0 N'$, and set $F_\beta(z) = f[\beta e(z)]$. $F_\beta(z)$ is defined on D' and analytic in its interior. Consequently,

$$f_2'(\beta) = \frac{\partial}{\partial y} F_\beta(z)|_{z=0} = iF_\beta'(0) = \frac{1}{2\pi}\int_C F_\beta(z)\frac{dz}{z^2} = \frac{1}{2\pi}\int_C f[\beta e(z)]\frac{dz}{z^2}.$$

However, $[\beta e(z)]\,[\alpha_0 e(z)]^{-1}$ is in N, for z on C, so that

$$|f[\beta e(z)] - f[\alpha_0 e(z)]| < \epsilon, \text{ all } z \in C.$$

Thus

$$|f_2'(\beta) - f_2'(\alpha_0)| \leq \frac{\epsilon}{2\pi}\int_C \frac{|dz|}{|z|^2},$$

where

$$\int_C \frac{|dz|}{|z|^2}$$

is a fixed constant. This proves the continuity of f_2'; from $f_1' = -if_2'$ follows the continuity of f_1'.

A function ρ on G will be called *differentiable* at $u \in G$, if $(d/dy)\ [uh(e^y)]\big|_{y=0}$ exists; and, if so, we shall denote the derivative by $\rho'(u)$. A continuous, almost everywhere continuously differentiable function ρ from G to R defines a *curve* in M, as the set $[\rho(u),u]$, $u \in G$. If f is a function defined on C, we shall set

$$\int_C f(\alpha)\ d\alpha = \int_G f[\rho(u),\ u]\left[1 - i\,\frac{\rho'(u)}{\rho(u)}\right] du,$$

the integral on the right being taken with respect to the Haar measure in G. (A curve as defined here is the analogue of the simple closed curve in the plane which encloses the origin.)

THEOREM 2. *Let f be analytic in an annulus, $a < r < b$, $u \in G$, and C_1, C_2, be two curves lying within the annulus. Then*

$$\int_{C_1} f(\alpha)\ d\alpha = \int_{C_2} f(\alpha)\ d\alpha.$$

The proof of this result is elementary and will be omitted.

Suppose that f is analytic in an annulus, $a < r < b$, and C_1 and C_2 are two curves defined by functions ρ_1 and ρ_2 on G, for which $a < \rho_1(u) < \rho_2(u) < b$. Suppose, furthermore, that f does not vanish on either C_1 or C_2. Then

$$N(C_1,\ C_2) = \int_{C_2} \frac{f_1'(\alpha)}{f(\alpha)} d\alpha - \int_{C_1} \frac{f_1'(\alpha)}{f(\alpha)}\ d\alpha$$

exists and may be interpreted as giving a measure of the zeros of f between C_1 and C_2 (although not simply the number of zeros). This idea will be developed in detail in a later note.

POLYADIC BOOLEAN ALGEBRAS

BY PAUL R. HALMOS

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CHICAGO

Communicated by Saunders Mac Lane, March 25, 1954

1. *Introduction.*—It is well known that when a logician speaks of the propositional calculus and when a mathematician speaks of Boolean algebras there is a sense in which they are talking about the same thing. What, in that sense, is the mathematical version of the so-called first-order functional calculus? Several answers to this question have been proposed in recent years;[1] the most satisfying of these answers is the concept of cylindric algebra introduced by Tarski and Thompson. The answer offered in this note (namely, polyadic Boolean algebra) differs

from previous proposals mainly in being able to imitate the substitution processes that give the functional calculi their characteristic flavor. The imitation appears to be successful. It turns out, for example, just as it does in the propositional case, that the customary approach to the functional calculus can be viewed as nothing but a detailed description of the symbolism of a *free* polyadic algebra.

It should be emphasized that the theory of polyadic algebras is discussed here not as a possible tool for solving problems about the foundations of mathematics, but as an independently interesting part of modern algebra. The possibility of making a dictionary that translates logical terms (such as *interpretation* and *semantic completeness*) into algebraic terms (such as *homomorphism* and *semisimplicity*) is to be regarded as merely an amusing by-product of the discussion.

2. *Monadic Algebras.*—If a *proposition* is taken to mean an element of a Boolean algebra $\boldsymbol{B}$, then it is natural to interpret a *propositional function* as a function from some set X into $\boldsymbol{B}$. The set $\boldsymbol{A}^*$ of all such functions is a Boolean algebra with respect to the pointwise operations. A *functional monadic algebra* is a Boolean subalgebra $\boldsymbol{A}$ of $\boldsymbol{A}^*$ such that, for each p in $\boldsymbol{A}$, the range of p has a supremum in $\boldsymbol{B}$, and such that the (constant) function $\exists p$, whose value at every point of X is that supremum, belongs to $\boldsymbol{A}$. The mapping $\exists$ of $\boldsymbol{A}$ into itself is called a *functional quantifier;* it is easy to verify that it satisfies the conditions

$$(Q_1) \quad \exists 0 = 0,$$
$$(Q_2) \quad p \leqq \exists p,$$
$$(Q_3) \quad \exists(p \wedge \exists q) = \exists p \wedge \exists q,$$

for all p and q in $\boldsymbol{A}$. (Properly speaking, $\exists$ should be called a functional *existential* quantifier; *universal* quantifiers are defined by an obvious dualization.) A general concept that applies to any Boolean algebra is obtained by abstraction from the functional case: a *quantifier* is a mapping $\exists$ of a Boolean algebra into itself satisfying (Q_1)–(Q_3). (This concept occurs also in the announcement of the results of Tarski and Thompson.) A *monadic algebra* is a Boolean algebra with a quantifier. The entire theory of Aristotelian (syllogistic) logic is easily subsumed under the elementary algebraic theory of monadic algebras.

3. *Polyadic Algebras.*—Propositional functions of several variables are treated by considering functions from a Cartesian power, say X^I, into a Boolean algebra $\boldsymbol{B}$. Even if the index set I is infinite, the interesting functions are those that depend on finitely many co-ordinates only. The set $\boldsymbol{A}^*$ of all such functions is a Boolean algebra, as before.

Let T^* be the semigroup of all transformations of I into itself (not necessarily one-to-one and not necessarily onto). Every element τ of T^* induces a Boolean endomorphism $p \to \tau p$ on $\boldsymbol{A}^*$, defined by allowing τ to act on the indices of the co-ordinates of the argument of p. [Example: if $I = \{1, 2, 3\}$ and if τ is the cyclic permutation $(1, 2, 3)$, then $\tau p(x_1, x_2, x_3) = p(x_2, x_3, x_1)$.]

If $x \in X^I$ and $J \subset I$, it is natural to consider the set of points in X^I that can be obtained by varying arbitrarily those co-ordinates of x whose index is in J. [Example: if $I = \{1, 2, 3\}$, $x = (x_1, x_2, x_3)$, and $J = \{1, 3\}$, then the set so obtained consists of all points whose second co-ordinate is x_2.] If $p \in \boldsymbol{A}^*$ and if the image of that set under p has a supremum in $\boldsymbol{B}$, that supremum is denoted by $\exists_J p(x)$; if J contains only one element, say i, it is convenient to write $\exists_i$ for $\exists_J$.

A *functional polyadic algebra* (in more detail, a $\boldsymbol{B}$-valued functional polyadic algebra with domain X) is a Boolean subalgebra $\boldsymbol{A}$ of $\boldsymbol{A}^*$ such that, for each p in $\boldsymbol{A}$, (1) $\exists_J p(x)$ exists whenever $J \subset I$ and $x \in X^I$, (2) the function $\exists_J p$ belongs to $\boldsymbol{A}$ whenever $J \subset I$, and (3) $\tau p \in A$ whenever $\tau \in T^*$. Because of the finiteness condition imposed on the elements of $\boldsymbol{A}^*$, it turns out that it is sufficient to assume (1) and (2) for finite subsets J and to assume (3) for finite transformations τ; the infinite ones can then be recaptured. (A transformation in T^* is *finite* if it coincides with the identity transformation outside some finite subset of I. The set T of all finite transformations is a subsemigroup of T^*.)

The operators $\exists_J$ are quantifiers on $\boldsymbol{A}$; the quantifiers $\exists_J$ (J finite) and the endomorphisms τ(τ finite) have the following properties.

(P_1) *If J is empty, then $\exists_J p = p$ for all p.*

(P_2) $\exists_J \exists_K = \exists_{J \cup K}$.

(P_3) *If τ is the identity, then $\tau p = p$ for all p.*

(P_4) $(\sigma\tau)p = \sigma(\tau p)$.

(P_5) *If $\sigma = \tau$ outside J, then $\sigma \exists_J = \tau \exists_J$.*

(P_6) *If τ is one-to-one on $\tau^{-1}J$, then $\exists_J \tau = \tau \exists_{\tau^{-1}J}$.*

(P_7) *For every p in $\boldsymbol{A}$ there exists a finite subset J of I such that $\exists_i p = p$ whenever i is not in J.*

(The transformation τ is said to be one-to-one on $\tau^{-1}J$ if it never maps two distinct elements of I on the same element of J.)

The superficially complicated results (P_5) and (P_6) are merely a telegraphic summary of the usual intuitively obvious relations between quantification and substitution. A couple of examples will make them clear. Suppose that i and j are distinct elements of I, and let τ be the transformation that maps i on j and maps everything else (including j) on itself. Since τ agrees with the identity outside $\{i\}$, it follows from (P_5) that $\tau \exists_i = \exists_i$. This relation corresponds to the familiar fact that once a variable has been quantified, the replacement of that variable by another one has no further effect. To get another example, note, for the same τ, that $\tau^{-1}\{i\}$ is empty; it follows from (P_6) and (P_1) that $\exists_i \tau = \tau$. This relation corresponds to the familiar fact that once a variable has been replaced by another one, a quantification on the replaced variable has no further effect.

Once again a general concept is obtained by abstraction. A *polyadic algebra* is a Boolean algebra $\boldsymbol{A}$, an index set I (whose elements are called *variables*), a correspondence from finite subsets of I to quantifiers on $\boldsymbol{A}$, and a correspondence from finite transformations on I to Boolean endomorphisms on $\boldsymbol{A}$, such that (P_1)–(P_7) are satisfied. It is useful to know that, just as in the functional case, the infinite quantifiers and endomorphisms can be defined in terms of the finite ones, and these extended operators satisfy the same conditions as their finite counterparts.

4. *Polyadic Logics.*—The usual machinery of universal algebra is immediately applicable to polyadic Boolean algebras; terms such as *subalgebra*, *homomorphism*, *kernel*, *ideal*, *maximal ideal*, *simple algebra*, and *free algebra* are defined as always. The essential point is to insist, in every case, that the corresponding Boolean concepts be invariant under the operators (quantifiers and endomorphisms) that define the polyadic structure. (Caution: polyadic algebras with different sets of variables are distinct types of algebraic systems. A homomorphism, for instance,

is defined only between algebras with the same sets of variables. The situation is analogous to the theory of groups with operators.) A minor difficulty occurs in the definition of a free algebra: because of the assumed finiteness condition (P_7), a polyadic algebra can never be really free. The difficulty is easily conquered by an appropriate relativization of the concept of freedom.

With the algebraic nomenclature at hand, many logical terms are easily translatable into algebraic language. A *polyadic logic*, to begin with, is defined as a pair $(\boldsymbol{A}, \boldsymbol{I})$, where $\boldsymbol{A}$ is a polyadic algebra and $\boldsymbol{I}$ is a polyadic ideal in $\boldsymbol{A}$. The motivation for this definition is the logical concept of refutability. In the customary treatment of logic, certain elements of an appropriate algebraic system (usually, though only implicitly, a polyadic algebra) are called refutable. From the algebraic point of view the definition of refutability in any particular case is inessential. What is important is the structure of the set $\boldsymbol{I}$ of all refutable elements; it turns out (and intuitive considerations demand that it should turn out so) that $\boldsymbol{I}$ is a polyadic ideal. (This discussion is, in fact, the dual of the usual one; reasons of algebraic convenience make refutability a more desirable concept than provability.)

A polyadic logic $(\boldsymbol{A}, \boldsymbol{I})$ is *simply consistent* if $\boldsymbol{I}$ is a proper ideal in $\boldsymbol{A}$ (not everything is refutable). Simple consistency is equivalent to the requirement that for no element p of $\boldsymbol{A}$ do both p and p' belong to $\boldsymbol{I}$ (where p' is the Boolean complement, or negation, of p). A polyadic logic $(\boldsymbol{A}, \boldsymbol{I})$ is *simply complete* if either $\boldsymbol{I} = \boldsymbol{A}$ or else $\boldsymbol{I}$ is a maximal ideal in $\boldsymbol{A}$. Simple completeness is equivalent to the requirement that, for every closed element p of $\boldsymbol{A}$, either p or p' belongs to $\boldsymbol{I}$. (An element p of $\boldsymbol{A}$ is *closed* if $\exists_I p = p$.) In these terms the celebrated Gödel incompleteness theorem asserts that certain important polyadic logics are either incomplete or inconsistent.

5. *Semantic Completeness.*—Since every simply consistent polyadic logic $(\boldsymbol{A}, \boldsymbol{I})$ is easily studied in terms of the associated polyadic quotient algebra $\boldsymbol{A}/\boldsymbol{I}$, there is no essential loss of generality in restricting attention to algebras instead of logics. It is, of course, necessary to keep in mind that the reduction to algebras reduces some non-trivial logical concepts to trivialities; thus, for instance, the only "refutable" element of $\boldsymbol{A}/\boldsymbol{I}$ is the zero element.

"Semantic" concepts refer to the representation theory of polyadic algebras. (The intrinsic, or structural, concepts of the preceding section are usually called "syntactic.") Representation theory proceeds, as always, by selecting a class of particularly simple and "concrete" polyadic algebras and asking to what extent every algebra is representable in terms of algebras of that class. For intuitively obvious reasons the logically important "concrete" algebras are the two-valued functional algebras, i.e., the functional algebras whose elements are functions from a Cartesian power X^I into the two-element Boolean algebra. Such algebras are called *models*.

An *interpretation* of an algebra $\boldsymbol{A}$ in a model $\boldsymbol{B}$ is a polyadic homomorphism from $\boldsymbol{A}$ onto $\boldsymbol{B}$. An element p of an algebra $\boldsymbol{A}$ is *universally invalid* if it is "false" in every interpretation, i.e., in algebraic language, if $\varphi p = 0$ whenever φ is a homomorphism from $\boldsymbol{A}$ on a model. (Once again, the present discussion is the dual of the usual one.) The algebra $\boldsymbol{A}$ is *semantically complete* if the only universally invalid element is 0. In logical terms, $\boldsymbol{A}$ is semantically complete if every universally invalid element is refutable, or, dually, if every universally valid element is provable.

6. *Constants.*—An important concept in the representation theory of polyadic

algebras is that of a constant. Intuitively, a constant is something that can be substituted for a variable; in the functional case a typical example of a constant is an element of the domain. The algebraic essence of the concept is captured by examining (either in terms of intuitive logic or in a functional example) the effect of replacing some of the variables (say the ones corresponding to a finite subset J of I) by an intuitively conceived constant. Abstraction from these special cases suggests the following definition. A *constant* (of a polyadic algebra $\boldsymbol{A}$ with variables $\boldsymbol{I}$) is a mapping $J \to c(J)$ from finite subsets of $\boldsymbol{I}$ to Boolean endomorphisms of $\boldsymbol{A}$ such that:

(C_1) *If J is empty, then $c(J)p = p$ for all p,*
(C_2) $c(J)\,c(K) = c(J \cup K)$,
(C_3) $c(J)\,\exists_K = \exists_K\, c(J - K)$,
(C_4) $\exists_J\, c(K) = c(K)\,\exists_{J-K}$,
(C_5) $c(J)\,\tau = \tau c(\tau^{-1}J)$ *for every finite transformation τ on I.*

There are simple examples of polyadic algebras that do not possess any constants. A polyadic algebra $\boldsymbol{A}$ with variables I is *rich* if, whenever $p \in \boldsymbol{A}$ and $i \in I$, there exists a constant c such that $\exists_i p = c(i)p$. (In highly informal language: there is a witness to every existential proposition.) The most important fact about rich algebras is that there are enough of them.

THEOREM 1. *Every polyadic algebra is isomorphic to a subalgebra of a rich algebra.*

The idea of the proof of Theorem 1 is most easily illustrated by sketching the proof of a special case. Suppose that $\boldsymbol{A}$ is a polyadic algebra with variables $\boldsymbol{I}$, and suppose that $p_0 \in \boldsymbol{A}$ and $i_0 \in I$. Let I^+ be a set containing one more element than $\boldsymbol{I}$, and consider a free polyadic algebra, with variables I^+, generated by the elements of $\boldsymbol{A}$. If that free algebra is reduced modulo all the relations that hold among the elements of $\boldsymbol{A}$, the result is a polyadic algebra $\boldsymbol{A}^+$ with variables I^+ (and hence, *a fortiori*, with variables I) that includes $\boldsymbol{A}$. If, for every finite subset J of I, $c(J)$ is defined to be the Boolean endomorphism of $\boldsymbol{A}^+$ induced by the transformation on I^+ that replaces the elements of J by the new element of I^+ and is otherwise equal to the identity, then c is a constant of the algebra $\boldsymbol{A}^+$ regarded as having the variables of I only. One more reduction, modulo the relation $\exists_{i_0} p_0 = c(i_0)p_0$, yields a polyadic extension of $\boldsymbol{A}$ that is rich enough, at least, to supply a witness for $\exists_{i_0} p_0$.

7. *Semisimplicity.*—The simplest representation theorem for polyadic algebras follows easily from an algebraic adaptation of the method of Rasiowa and Sikorski;[2] an analogue of this result for cylindric algebras was announced by Tarski.[3]

THEOREM 2. *Every polyadic algebra with an infinite set of variables is isomorphic to a functional algebra.*

Unfortunately Theorem 2 does not yield enough information about the structure of a general polyadic algebra. To state the desired sharper result, it is convenient to introduce the functional counterpart of the concept of richness. A functional algebra $\boldsymbol{A}$ with variables I and domain X is *functionally rich* if, whenever $p \in \boldsymbol{A}$, $i \in I$, and $x \in X^I$, there exists a point y of X^I such that all the co-ordinates of y, with the possible exception of the i co-ordinate, are the same as the corresponding co-ordinates of x and such that $\exists_i p(x) = p(y)$.

THEOREM 3. *Every polyadic algebra $\boldsymbol{A}$ with variables I is isomorphic to a subalgebra*

of a rich functional algebra the cardinal number of whose domain is less than or equal to the sum of the cardinal numbers of $\boldsymbol{A}$ *and of* I.

The proof of Theorem 3 is based on Theorem 1; the method is a modification of the one introduced by Henkin.[4] To construct the promised functional algebra, it is necessary to construct a domain X and a value algebra $\boldsymbol{B}$. The role of X is played by a sufficiently large set of constants for a rich polyadic extension of the given algebra $\boldsymbol{A}$, and the role of $\boldsymbol{B}$ is played by $\boldsymbol{A}$ itself.

The algebraic analogues of the Skolem-Löwenheim theorem and the Gödel completeness theorem are relatively easy consequences of Theorem 3. To discuss the latter, for instance, consider a functionally rich $\boldsymbol{B}$-valued functional algebra, and let φ_0 be an arbitrary Boolean homomorphism from $\boldsymbol{B}$ to the two-element Boolean algebra. The assumption of richness implies that the application of φ_0 to the values of the elements of the functional algebra induces a polyadic homomorphism φ into a two-valued functional algebra; more precisely, φ is defined by $\varphi p(x) = \varphi_0[p(x)]$. This technique, combined with Theorem 3, yields the following basic result.

THEOREM 4. *Every simple polyadic algebra is isomorphic to a model.*

The method of Rasiowa and Sikorski also serves to prove a theorem of this type, and it is in some respects simpler than Henkin's; its disadvantage is that it works under severe infinity and countability restrictions only.

The converse of Theorem 4 is true, easy, and, incidentally, independent of the preceding relatively high-powered representation theorems.

THEOREM 5. *Every model is simple.*

Since, on universal algebraic grounds, the kernel of a polyadic homomorphism is a maximal ideal if and only if its range is simple, semantic completeness is equivalent to the requirement that the intersection of all maximal ideals is $\{0\}$. Since, in analogy with other parts of algebra, it is natural to call a polyadic algebra satisfying this condition *semisimple*, it follows that a polyadic algebra is semantically complete if and only if it is semisimple. It is now easy to state the algebraic version of the Gödel completeness theorem.

THEOREM 6. *Every polyadic algebra is semisimple.*

The assertion of Theorem 6 for a polyadic algebra with the empty set of variables (i.e., for an ordinary Boolean algebra) is the most important part of Stone's theorem on the representation of Boolean algebras. Theorem 6 is, in fact, a relatively easy consequence of Stone's theorem; it is much nearer to the surface than, for instance, Theorem 3. In other words, the hard thing in proving Gödel's completeness theorem is not to prove semisimplicity but to prove the appropriate representation theorem (Theorem 4) that guarantees that semisimplicity is the right thing to prove.

[1] A. Tarski, *J. Symbolic Logic*, **6**, 73–89 (1941); C. J. Everett and S. Ulam, *Amer. J. Math.*, **68**, 77–88 (1946); F. I. Mautner, *Amer. J. Math.*, **68**, 345–384 (1946); A. Tarski and F. B. Thompson, *Bull. Amer. Math. Soc.*, **58**, 65 (1952).

[2] H. Rasiowa and R. Sikorski, *Fund. Math.*, **37**, 193–200 (1950).

[3] A. Tarski, *Bull. Amer. Math. Soc.*, **58**, 65 (1952).

[4] L. Henkin, *J. Symbolic Logic*, **14**, 159–166 (1949)

STUDIES IN THE CONFORMAL MAPPING OF RIEMANN SURFACES. II

By Maurice Heins

BROWN UNIVERSITY

Communicated by Marston Morse, March 4, 1954

1. The object of the present note is to announce further results of the investigation summarized partially in "Studies . . . I" (these Proceedings, **39,** No. 4, 322–324 [April, 1953]). We shall be concerned with (1) a property of the quasi-bounded component v_q of the residual term u_q; (2) an analysis of the notion of an asymptotic point in which we discern a subclass of such points distinguished by metric properties; and (3) maps locally of type-B*l*. "Theorem A" and "Lemma C" refer to "Studies . . . I."

2. *A Property of v_q.*—In the setting of Theorem A we have:

Theorem 2.1. *For each $p \in F$, the function $q \to v_q(p)$ is a bounded Green's potential on G.*

This theorem stands in interesting relation to certain notions introduced by Lehto.[1] Consider now two arbitrary Riemann surfaces F and G, the first being noncompact, no restrictions being placed on the ideal boundaries. Let f denote a directly conformal map of F into G, p_0 a given point of F, and Ω a relatively compact region of F containing p_0. If the frontier of Ω consists of a finite number of disjoint closed analytic Jordan curves, then it is easy to show that

$$\sum_{f(p) = q,\ p,\ \in \Omega} n(p)\mathcal{G}_\Omega(p,\ p_0)$$

is subharmonic and continuous on $G - \{f(p_0)\}$.[1] Consider the unrestricted class of Ω and introduce

$$S_{p_0}(q) = \sup_{\Omega} \Big\{ \sum_{f(p) = q,\ p \in \Omega} n(p)\mathcal{G}_\Omega(p,\ p_0) \Big\}. \tag{2.1}$$

Then S_{p_0} is lower semicontinuous on G. Let $S_{p_0}{}^*$ denote the upper-limit function of S_{p_0}. We have:

Theorem 2.2. *If G has positive boundary, then $S_{p_0}(q) = \sum_{f(p) = q} n(p)\mathcal{G}_F(p,\ p_0)$ and $S_{p_0}{}^*(q) = \mathcal{G}_G[f(p_0),q] - v_q(p_0)$.*

3. *Mappings of Nonnegative Harmonic Functions.*—Let P denote the class of nonnegative harmonic functions with domain F. Let $\Omega(\subsetneqq F)$ denote a region of F which is not relatively compact and has the property that each point of $fr\Omega$ is in a continuum contained in $fr\Omega$. Let P_Ω denote the class of nonnegative harmonic functions with domain Ω which vanish continuously on $fr\Omega$. We introduce a map λ_Ω of P into P_Ω defined by the requirement that for each $u \in P$, $\lambda_\Omega(u)$ be the largest member of P_Ω dominated by u. The mapping λ_Ω is additive and homogeneous. For each $U \in P_\Omega$, let U^* be defined on F as equal to U on Ω and equal to 0 on $F - \Omega$, so that U^* is subharmonic on F. Let Q_Ω denote the subclass of P_Ω consisting of those U with U^* possessing a finite harmonic majorant, and let μ_Ω denote the map of Q_Ω into P which maps $U \in Q_\Omega$ into the least harmonic majorant of U^*. The mapping μ_Ω also is additive and homogeneous. Special cases of such maps and related ones have been considered by Parreau[2] and the author.[3]

We recall that $u(\not\equiv 0) \in P(P_\Omega)$ is termed *minimal*, provided that for all $v \in P(P_\Omega)$ which satisfy $v \leqslant u$, v is proportional to u.[4]

Fundamental properties of λ_Ω and μ_Ω are given by:

THEOREM 3.1. *(a) $\lambda_\Omega o \mu_\Omega$ is the identity map of Q_Ω onto itself. (b) Given $U \in Q_\Omega$, U is respectively quasi-bounded, singular, minimal (on Ω) if and only if $\mu_\Omega(U)$ is respectively quasi-bounded, singular, minimal (on F). (c) If $\{\Omega_k\}$ is a family of Ω satisfying $\Omega_k \cap \Omega_l = \varnothing$ for $k \neq l$, $v \in P$, and $U_k (\in Q_{\Omega_k})$ satisfies $\mu_{\Omega_k}(U_k) \leqslant v$ for each k, then $\Sigma\mu_{\Omega_k}(U_l) \leqslant v$.*

This theorem leads to the following localization principle: In the situation considered in Theorem A, let ω denote a relatively compact region of G whose frontier consists of a finite number of disjoint Jordan curves; let $q \epsilon \omega$; let $\{\Omega_k\}$ denote the set of components of $f^{-1}(\omega)$ which are not relatively compact ($\Omega_k \cap \Omega_l = \varnothing$ for $k \neq l$); let $u_q^{(k)}$ denote the greatest harmonic minorant of $\mathcal{G}_\omega[f_{\Omega_k}(p), q]$, where f_{Ω_k} is the restriction of f to Ω_k; and let $w_q(w_q^{(k)})$ denote the singular component of $u_q(u_q^{(k)})$. Then

$$0 \leqslant u_q - \sum \mu_{\Omega_k}(u_q^{(k)}) = 0(1); \tag{3.1}$$

$$w_q = \sum \mu_{\Omega_k}(w_q^{(k)}).$$

The obvious interpretations are to be made in the cases where $f^{-1}(\omega) = F$ or $\{\Omega_k\}$ is empty.

4. *Asymptotic Spots.*—We continue in the setting of Theorem A. We introduce the notion of an asymptotic spot as follows: Let $q \in G$ and let Φ denote the family of simply connected Jordan regions containing q. By an *asymptotic spot* over q is meant a function σ with domain Φ satisfying the following conditions: (1) for each $\omega \in \Phi$, $\sigma(\omega)$ is a component of $f^{-1}(\omega)$ which is not relatively compact; (2) if $\omega_1 \subset \omega_2$, then $\sigma(\omega_1) \subset \sigma(\omega_2)$. This definition is inspired by the classical function-theoretic notion of a tract. We term a point $q \in G$ an *asymptotic point* of f, provided that there exists an asymptotic spot over q. This definition is equivalent to the standard definition in terms of paths.

It is known (Mazurkiewicz, Kierst) that the set of asymptotic points of f is an analytic set in G. On the other hand, given G and an analytic set $E \subset G$, there exists a directly conformal map f of $|z| < 1$ into G whose set of asymptotic points is E. (These statements hold without restrictions on F and G.)

Our objective is to discern an important subclass of asymptotic spots distinguished by metric properties. We consider, for an asymptotic spot σ over q, the family

$$\{\mu_{\sigma(\omega)}[U_{\sigma(\omega)}]\},$$

where $u_{\sigma(\omega)}$ is the greatest harmonic minorant of $\mathcal{G}_\omega[f_{\sigma(\omega)}(p), q]$, $f_{\sigma(\omega)}$ being the restriction of f to $\sigma(\omega)$. There exists a unique maximal $u_\sigma \in P$ dominated by $\mu_{\sigma(\omega)}[u_{\sigma(\omega)}]$ for all $\omega \in \Phi$.

If $u_\sigma \not\equiv 0$, then we term σ a *metric asymptotic spot* and observe that u_σ is singular on F and is dominated by w_q. We term a point q of G a *metric asymptotic point* of f, provided that there exists a metric asymptotic spot over q. We summarize our results concerning metric asymptotic spots and points in:

THEOREM 4.1. *(1) The set of metric asymptotic spots of f over a given point $q \in G$ is countable. (2) The set of metric asymptotic points of f is an F_σ in G of capacity zero (and such a set is always realizable as the set of metric asymptotic points for a suitably chosen conformal mapping into G). (3) If σ is logarithmic,[5] then σ is a metric asymp-*

totic spot and u_σ is minimal. (4) *If F has finite genus and finite contour number and σ is a metric asymptotic spot, then u_σ is minimal.*

We remark that (4) fails to hold if the restriction on F is omitted. A simple example is furnished by considering $z \sin z$ on $|z \sin z| > 1$ as a map of this region into $|w| > 1$ and taking $q = \infty$.

5. *Properties of Maps Which Are Locally of Type-Bl.*—From this point on, it will be convenient to agree that the Riemann surfaces considered are unrestricted save explicit mention to the contrary. We introduce the valence function, ν_f, for f defined by

$$\nu_f(q) = \sum_{f(p) = q} n(p),\ q \in G.$$

If $\nu_f \equiv$ const. $< +\infty$, then f is locally of type-Bl. In the other direction one may show for maps of type-Bl:

THEOREM 5.1. *If for some $q_0 \in G$, $u_{q_0} = 0$ and $\nu_f(q_0) < +\infty$, then $\nu_f(q) \leqq \nu_f(q_0)$ for all $q \in G$ and the set of q for which $\nu_f(q) < \nu_f(q_0)$ is a closed set of capacity zero. This set is the set of q for which $u_q > 0$.*

A consequence of Theorem 5.1 is that, if f is a directly conformal map of F into G (F and G unrestricted) which is locally of type-Bl, then either $\nu_f = +\infty$, except for a set of capacity zero, or else $\nu_f = c$ (const.) $< +\infty$, except for a closed set of capacity zero on which $\nu_f < c$.

6. *Transitivity Properties.*—Consider, now, Riemann surfaces F, G, H, and directly conformal maps f of F into G and g of G into H, and let $h = g \circ f$. Then a necessary and sufficient condition that h be locally of type-Bl (relative to H) is that f be locally of type-Bl (relative to G) and that g be locally of type-Bl (relative to H).

7. *Sufficient Conditions That a Map Be Locally of Type-Bl.*—We consider, now, Riemann surfaces F and G and a directly conformal map f of F into G. The set of points $E \subset G$ at which f is not of type-Bl is closed. With the aid of Lemma C, one is led to the following sharper result: If ω is an open set of G such that $E \cap \omega \neq \varnothing$, then $E \cap \omega$ has positive capacity.

This result leads to the following sufficient condition that a map be locally of type-Bl: If F does not admit nonconstant bounded harmonic functions, then f is locally of type-Bl. We remark in passing that the existence of such F with positive boundary has been established by Y. Tôki.[6] Further, we note that, if F has positive boundary and G null boundary and f is locally of type-Bl, then $\nu_f(q) = +\infty$ except for a set of capacity zero.

A second sufficient condition for a map to be locally of type-Bl may be given in terms of the set $\mathfrak{A}$ of its asymptotic points. In fact, f is locally of type-Bl, if $\mathfrak{A}$ has zero (inner) capacity.

The parentheses call for a gloss. We observe that the fact that $\mathfrak{A}$ is analytic implies, by a result of G. Choquet,[7] that $\mathfrak{A}$ is capacitable. However, the proof of our theorem is based upon Frostman's generalization of the theorem of F. and M. Riesz and appeals only to the fact that the inner capacity of $\mathfrak{A}$ is zero.

[1] O. Lehto, "A Majorant Principle in the Theory of Functions," *Math. Scandinav.*, **1**, 5–17 (1953).

[2] M. Parreau, "Sur les moyennes des fonctions harmoniques et analytiques et la classification des surfaces de Riemann," *Ann. Inst. Fourier* (Grenoble) (Paris thesis, 1952).

[3] M. Heins, "Riemann Surfaces of Infinite Genus," *Ann. Math.*, **55**, No. 2, 296–317 (1952).

[4] R. S. Martin, "Minimal Positive Harmonic Functions," *Trans. A.M.S.*, **49**, 137–172 (1941).

[5] I.e., there exists $\omega \in \Phi$ with the property that $\sigma(\omega)$ is simply connected and the restriction of f to $\sigma(\omega)$ is topologically equivalent to e^z restricted to $\Re z < 0$ (q corresponding to 0).

[6] Y. Tôki, "On the Classification of Open Riemann Surfaces," *Osaka Math. J.*, **4**, 191–202 (1952).

[7] G. Choquet, "Capacitabilité, théorèmes fondamentaux," *Compt. Rend. Acad. Sci.*, **234**, 784–786 (1952).

ON THE LIE RING OF A SIMPLE RING

BY I. N. HERSTEIN

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF PENNSYLVANIA

Communicated by Saunders Mac Lane, March 8, 1954

Given any associative ring A, one can render it into a Lie ring by defining a new product, the Lie product, for any two elements a and b in A by means of $[a, b] = ab - ba$, where ab is the old associative product in A. It is natural to expect that the Lie ring so obtained has a structure which is closely connected with the associative structure of A.

We consider the situation where A is a simple associative ring. A subset U of A is said to be a Lie ideal of A if it is an additive subgroup of A and if, in addition, whenever $u \in U$ then $xu - ux \in U$ for all $x \in A$. If $[A,A]$ is the additive subgroup of A generated by all the commutators $xy - yx$, then $[A,A]$ is a Lie ring under the Lie product of A. It has been conjectured that if A is of characteristic $\neq 2$, then any proper Lie ideal of $[A,A]$ must be in the center of A. In case A possesses two nontrivial orthogonal idempotents whose sum is not 1, then Kaplansky has shown the conjecture to hold true.[1]

We prove the following about the Lie ideal structure of A itself:

THEOREM 1. *If A is a simple, non-nil ring of characteristic $\neq 2$ and U is a Lie ideal of A, then either U is contained in the center of A, or else $U \supset [A,A]$.*

For characteristic 2 the result is false, as is shown by the example of a simple algebra of dimension 4 over its center, which is a field of characteristic 2. The next theorem shows that this constitutes the only counterexample to the theorem, namely:

THEOREM 2. *If A is a simple, non-nil ring of characteristic 2 and if A is not 4-dimensional over its center, then any Lie ideal U of A is either contained in the center of A or $U \supset [A, A]$.*

Of key importance in the proof of the above theorems and in that of the subsequent ones is the study of the associated subring, $S(U)$, of U, which is defined by $S(U) = \{x \in A \mid [x, A] \subset U\}$.

We turn to the conjecture concerning $[A,A]$. The following results are obtained:

THEOREM 3. *If A is a division ring of characteristic $\neq 2$, then any proper Lie ideal of $[A,A]$ is contained in the center of A.*

If A is not a division ring, by adding the additional proviso that A is not of characteristic 3, we can still prove the conjecture. That is:

THEOREM 4. *If A is a simple, non-nil ring of characteristic $\neq 2,3$, then any proper Lie ideal of $[A,A]$ is contained in the center of A.*

However, in characteristic 3 quite a bit of relevant information can be obtained; this information tends to affirm the validity of the conjecture even in characteristic 3. The results obtained are:

THEOREM 5. *If A is a simple ring of characteristic 3 and is not a nil-ring, then any proper Lie ideal of $[A,A]$ is solvable to a fixed degree k ($k = 12$ will do).*

From this it follows that:

THEOREM 6. *If A is a simple non-nil ring of characteristic 3 then $[[A,A],[A,A]] = [A,A]$.*

Of possible importance in disposing of the remaining case, that of characteristic 3, might be:

THEOREM 7. *If A is a simple, non-nil ring of characteristic 3, then $[A,A]$ possesses a unique, proper maximal Lie ideal U_M; if U is any proper Lie ideal of $[A,A]$, then $U \supset U_M$.*

If it should happen that A possessess only a trivial center, then again the conjecture is provable. This is expressed by:

THEOREM 8. *If A is a simple, non-nil ring of characteristic 3 and has (0) as its center, then $[A,A]$ is a simple Lie ring.*

If A is any associative ring, we can introduce another new product by defining $a \circ b = ab + ba$. This way we obtain the Jordan ring of A. An additive subgroup U of A is said to be a Jordan ideal if $u \in U$, $x \in A$ implies that $ux + xu \in U$. We also show that:

THEOREM 9. *If A is a simple associative ring of characteristic $\neq 2$ then A has no nontrivial Jordan ideals.*

Detailed proofs of these results and some other subsidiary theorems will appear elsewhere.

[1] Irving Kaplansky, unpublished paper.

ON A THEOREM BY MURNAGHAN

BY E. M. IBRAHIM

IBRAHIM UNIVERSITY, CAIRO

Communicated by F. D. Murnaghan, January 26, 1954

Murnaghan[1] mentioned that the irreducible representations $[\lambda]$ of the $2k$-dimensional orthogonal group and $\langle\lambda\rangle$ of the $2k$-dimensional unitary symplectic group are connected by means of the relation $\langle\lambda\rangle = [\lambda^*]^*$ (the * denotes the associated partition). If, then, (λ) is a partition of an even integer m,

$$\langle\lambda\rangle \otimes \{\mu\} = ([\lambda^*] \otimes \{\mu\})^*,$$

while if (λ) is a partition of odd integer m,

$$\langle\lambda\rangle \otimes \{\mu\} = ([\lambda^*] \otimes \{\mu^*\})^*.$$

These relations enabled Murnaghan to write down without any calculations the

analysis of $\langle\lambda\rangle \otimes \{\mu\}$ when the analyses of $[\lambda^*] \otimes \{\mu\}$ and of $[\lambda^*] \otimes \{\mu^*\}$ are known. The object of the present paper is to prove these theorems and to give other relations connecting the characters of the orthogonal and the symplectic groups.

It is known that the general orthogonal group in n variables is the group of transformations which leaves invariant an arbitrary nonsingular quadratic form $\sum g_{ij}x_ix_j$, while the symplectic group of transformations leaves the linear complex form of type $\{1^2\}$ invariant.

The characters of the orthogonal group and of the symplectic group have been found by Schur (1924) and Weyl (1939), respectively. The method used is transcendental and depends on integration over the group manifold. Later D. E. Littlewood[2] obtained these characters by purely algebraic methods and the following results have been established:

$$[\lambda] = \{\lambda\} + \sum(-)^{p/2}\Gamma_{\nu\mu\lambda}\{\mu\}, \tag{1}$$

where $\Gamma_{\nu\mu\lambda}$ is the coefficient of $\{\lambda\}$ in the product $\{\nu\}\ \{\mu\}$. $\{\nu\}$, a partition of p, is summed for all partitions which in Frobenius' nomenclature are of one of the forms

$$\binom{r+1}{r}, \binom{r+1 \quad s+1}{r \quad\ s}, \binom{r+1 \quad s+1 \quad t+1}{r \quad\ s \quad\ t} \ldots.$$

These partitions appear as

$$\begin{aligned} 1 - \{2\} + \{31\} - \{41^2\} - \{3^2\} - \{42^2\} + \{431\} + \{4^3\} + \\ \{51^3\} - \{531^2\} + \{5421\} - \{5^22^2\} - \{5^33\} + \\ \{\lambda\} = \langle\lambda\rangle + \sum\Gamma_{\beta\mu\lambda}\langle\mu\rangle, \end{aligned} \tag{2}$$

where $\Gamma_{\beta\mu\lambda}$ is the coefficient of $\{\lambda\}$ in the product $\{\beta\}\ \{\mu\}$, $\{\beta\}$ is summed for all partitions in which each part is repeated an even nunber of times, i.e.,

$$1 + \{1^2\} + \{2^2\} + \{1^4\} + \{3^2\} + \{2^21^2\} + \{1^6\} + \ldots.$$

A direct relation connecting the characters of the orthogonal group and those of the symplectic group could be deduced as

$$[\lambda] = \langle\lambda\rangle + \Gamma_{\eta\mu\lambda}\langle\mu\rangle, \tag{3}$$

where $\Gamma_{\eta\mu\lambda}$ is the coefficient of $\{\lambda\}$ in the product $\{\eta\}\ \{\mu\}$, $\{\eta\}$ is summed for all partitions given by the product $\{\eta\} = \{\nu\}\ \{\beta\} = (1 - \{2\} + \{31\} - \{41^2\} - \{3^2\} - \{42^2\} + \{431\} + \{4^3\} \ldots)(1 + \{1^2\} + \{2^2\} + \{3^2\} + \{4^2\} + \{5^2\} + \{6^2\} + \{7^2\} + \{8^2\} \ldots) = 1 - \{2\} + \{1^2\} + \{2^2\} - \{21^2\} - \{2^3\} \ldots.$

Example: For the $2k$-dimensional orthogonal group,

$$[42] = \langle 42\rangle - \langle 4\rangle - \langle 2^2\rangle + \langle 2\rangle.$$

In fact, for the $2k$-dimensional orthogonal group when (λ) is a 2-element partition, we have

$$(a)\ [mn] = \{mn\} - \sum_{j=0}^{2}\{m-j, n+j-2\} + \sum_{j=0}^{2}\{m-j-1, n+j-3\} - \{m-3, n-3\},$$

(b) $\{mn\} = \sum_{j=0}^{n} \{m-j, n-j\}$,

(c) $[mn] = \langle mn\rangle - \langle m-2, n\rangle - \langle m, n-2\rangle + \langle m-2, n-2\rangle$: $m \neq n$,

(d) $[mm] = \langle mm\rangle - \langle m, m-2\rangle + \langle m-1, m-1\rangle + \langle m-2, m-2\rangle$.

Also it is known that

$$\langle\lambda\rangle = \{\lambda\} + \sum(-)^{p/2}\Gamma_{\alpha\mu\lambda}\{\mu\}, \tag{4}$$

summed for all partitions of the set $\{\alpha\}$, a partition of p, which in Frobenius' nomenclature is of one of the forms

$$\begin{pmatrix} r \\ r+1 \end{pmatrix}, \begin{pmatrix} r & s \\ r+1 & s+1 \end{pmatrix}, \begin{pmatrix} r & s & t \\ r+1 & s+1 & t+1 \end{pmatrix}, \ldots \text{ i.e., } 1 - \{1^2\} + \{21^2\} - \{2^3\} + \{32^21\} + \ldots$$

such that $\{\lambda\}$ appear in the product $\{\alpha\}\{\mu\}$ with the coefficient $\Gamma_{\alpha\mu\lambda}$. But

$$\{\lambda\} = [\lambda] + \sum\Gamma_{\delta\mu\lambda}[\mu], \tag{5}$$

where $\Gamma_{\delta\mu\lambda}$ is the coefficient of $\{\lambda\}$ in the product $\{\delta\}\{\mu\}$ and the summation is taken with respect to all partitions $\{\delta\}$ into even parts only, i.e., $1 + \{2\} + \{4\} + \{2^2\} + \{6\} + \{42\} + \{8\} + \{62\} + \ldots$. Therefore, it could be deduced that

$$\langle\lambda\rangle = [\lambda] + \sum\Gamma_{\sigma\mu\lambda}[\mu], \tag{6}$$

where $\Gamma_{\sigma\mu\lambda}$ is the coefficient of $\{\lambda\}$ in the product $\{\sigma\}\{\mu\}$ and the summation is taken for all $\{\sigma\} = \{\delta\}\{\alpha\} = (1 + \{2\} + \{4\} + \{2^2\} + \{6\} + \{42\} + \{8\} + \{62\} + \{4^2\} + \{82\} + \{64\} + \{84\} + \{6^2\} + \{86\} + \{8^2\} + \ldots)(1 - \{1^2\} + \{21^2\} \ldots) = 1 - \{1^2\} + \{2\} - \{31\} + \{4\} - \{51\} + \{2^2\} - \{3^2\} + \{6\} - \{71\} + \{42\} - \{53\} + \{8\} + \{62\} - \{73\} + \{4^2\} - \{5^2\} + \{82\} + \{64\} - \{75\} + \{84\} + \{6^2\} - \{7^2\} + \{86\} + \{8^2\} \ldots$.

Example: For the $2k$-dimensional symplectic group,

$$\langle 42\rangle = [42] + [4] + [2^2] + [2] - [1^2] + [0].$$

As the set $\{\eta\} = \{\nu\}\{\beta\}$ is conjugate with the set $\{\sigma\} = \{\alpha\}\{\delta\}$, then if formula (3) gives $[\lambda] = \sum\langle\mu\rangle$, formula (6) will give

$$\langle\lambda^*\rangle = \sum[\mu^*], \tag{7}$$

and we can say that if, $\{\lambda\} = \sum[\mu]$, then $\{\lambda^*\} = \sum\langle\mu^*\rangle$.

Example: For the $2k$-dimensional symplectic group,

$$\langle 42\rangle = [42] + [4] + [2^2] + [2] - [1^2] + [0];$$

therefore, for the $2k$-dimensional orthogonal group $[2^21^2] = \langle 2^21^2\rangle + \langle 1^4\rangle + \langle 2^2\rangle + \langle 1^2\rangle - \langle 2\rangle + \langle 0\rangle$. Also if $\{31\} = \langle 31\rangle + \langle 2\rangle$, then $\{21^2\} = [21^2] + [1^2]$.

THEOREM. $\sum\langle\lambda\rangle = (\sum[\lambda^*])^*$.

By formula (7) the left hand side $= \sum\sum[\mu]$, where $\langle\lambda\rangle = \sum[\mu]$ and, therefore, $[\lambda^*] = \sum\langle\mu^*\rangle$. The right-hand side $= (\sum[\lambda^*])^* = (\sum\sum\langle\mu^*\rangle)^* = \sum\sum[\mu]$, which proves the theorem.

COROLLARY. $\sum[\mu] = (\sum\langle\mu^*\rangle)^*$.

Example: $\langle 3\rangle + \langle 21\rangle + \langle 2\rangle = ([1^3] + [21] + [1^2])^*$.

Now the theorem stated by Murnaghan could be deduced from the theorem of conjugates of Littlewood,[3] i.e., if $[\lambda] \otimes \{\mu\} = \sum\{\nu\}$, where λ is a partition of an integer m, then $\langle\lambda^*\rangle \otimes \{\mu\} = \sum\{\nu^*\}$ if m is even and $\langle\lambda^*\rangle \otimes \{\mu^*\} = \sum\{\nu^*\}$ if m is odd. For the $2k$-dimensional group

$$\langle 2\rangle \otimes \{2\} = \langle 4\rangle + \langle 2^2\rangle + \langle 1^2\rangle + \langle 0\rangle;\ \text{then, } [1^2] \otimes \{2\} = [1^4] + [2^2] + [2] + [0];\ \text{also if } [3] \otimes \{2\} = [6] + [42] + [4] + [2^2] + [2] + [0], \text{ then } \langle 1^3\rangle \otimes \{1^2\} = \langle 1^6\rangle + \langle 2^21^2\rangle + \langle 1^4\rangle + \langle 2^2\rangle + \langle 1^2\rangle + \langle 0\rangle.$$

COROLLARY.

(i) Since $\{\lambda\} \otimes (\{\mu_1\} + \{\mu_2\} + \dots) = \{\lambda\} \otimes \{\mu_1\} + \{\lambda\} \otimes \{\mu_2\} + \dots$, then, if $\{\lambda\} \otimes [\mu] = \sum[\nu]$, where λ is a partition of an integer m, $\{\lambda^*\} \otimes [\mu] = \sum\langle\nu^*\rangle$ if m is even, and $\{\lambda^*\} \otimes \langle\mu^*\rangle = \sum\langle\nu^*\rangle$ if m is odd.

(ii) If $[\lambda] \otimes [\mu] = \sum[\nu]$, λ is a partition of m, then $\langle\lambda^*\rangle \otimes [\mu] = \sum\langle\nu^*\rangle$ (m even) and $\langle\lambda^*\rangle \otimes \langle\mu^*\rangle = \sum\langle\nu^*\rangle$ (m odd).

(iii) If $\langle\lambda\rangle \otimes [\mu] = \sum\{\nu\}$, λ a partition of m, then $[\lambda^*] \otimes [\mu] = \sum\{\nu^*\}$ (m even) and $[\lambda^*] \otimes \langle\mu^*\rangle = \sum\{\nu^*\}$ (m odd).

[1] F. D. Murnaghan, PROC. NATL. ACAD. SCI., **38**, 966–973 (1952).

[2] D. E. Littlewood, *Phil. Trans. R. Soc.*, ser. A, No. 809, **239**, 387–417 (1944).

[3] D. E. Littlewood, *ibid.*, No. 807, pp. 305–65.

---

*ON THE ARITHMETIC NORMALITY OF THE GRASSMANN VARIETY**

BY JUN-ICHI IGUSA

DEPARTMENT OF MATHEMATICS, HARVARD UNIVERSITY, AND KYOTO UNIVERSITY, JAPAN

Communicated by Oscar Zariski, March 1, 1954

In this paper we shall prove, by a method which is valid for a universal domain of arbitrary characteristic, a theorem of Severi to the effect that every positive divisor on a Grassmann variety, which is canonically imbedded in a projective space, is a complete intersection by a hypersurface of the ambient space. The proof is a simple combination of formal lemmas due to Hodge, the well-known characterization of normality due to Muhly-Zariski, and the rationality of the Grassmann variety. Actually, Severi gave two proofs[1] of this assertion, neither of which can be applied immediately to the case of prime characteristic. We note that Severi's theorem and the first main theorem for the unimodular group[2] are substantially equivalent. Therefore, we have also proved this latter theorem in the case of prime characteristic. As a consequence, we shall generalize to the case of arbitrary characteristic a remark of Hodge to the effect that any Chow form is a homogeneous polynomial in the Plücker co-ordinates of the indeterminates of the form.

The subject of this paper was called to my attention by Professor Zariski, and I should like to thank him for doing so, and also for his interest shown to my proof.

1. *Arithmetic Normality.*—The Grassmann variety V is defined in the following way. We fix a projective space L^n of n dimension over the universal domain $\mathbf{K}$.[3] If L^r is a linear variety in L^n which is spanned by a set of $r+1$ independent generic points $x_0, x_1, \ldots, x_r$ of L^n over a field K, then L^r is defined over the purely transcendental extension $K(x_0, x_1, \ldots, x_r)$ of K. Therefore, its Chow point p has a locus V over K, which coincides with the variety of positive cycles of dimension r and of degree 1 in L^n. It is a simple matter to see that the homogeneous co-ordinates of the point p coincide with the Plücker co-ordinates

$$p_{i_0 i_1 \ldots i_r} = \begin{vmatrix} x_{0 i_0} x_{0 i_1} \ldots x_{0 i_r} \\ x_{1 i_0} x_{1 i_1} \ldots x_{1 i_r} \\ \ldots \quad \ldots \\ x_{r i_0} x_{r i_1} \ldots x_{r i_r} \end{vmatrix}$$

of L^r, where $x_{\alpha 0}, x_{\alpha 1}, \ldots, x_{\alpha n}$ are the homogeneous co-ordinates of x_α for $\alpha = 0, 1, \ldots, r$. Since the group of projective transformations in L^n operates transitively on the totality of L^r, we conclude that V is a "homeogeneous variety." Therefore, V is nonsingular and a fortiori everywhere locally normal.

We take a set of letters X_{ij} and denote by capitals $P_{i_0 i_1 \ldots i_r}$ the expressions which result if in the above determinants $p_{i_0 i_1 \ldots i_r}$ the x_{ij} are replaced by X_{ij}. Our purpose is to prove the following:

THEOREM 1 (ARITHMETIC NORMALITY). *The ring of polynomials* $\mathbf{K}[P]$ *in* $P_{i_0 i_1 \ldots i_r}$ *with coefficients in* $\mathbf{K}$ *is integrally closed.*

This can be done by a combination of some lemmas due to Hodge and stated below.

Any monomial of degree m

$$P_{i_0 i_1 \ldots i_r} P_{j_0 j_1 \ldots j_r} \ldots P_{k_0 k_1 \ldots k_r}$$

can be described by the following diagram

$$\begin{matrix} i_0 & j_0 & \ldots & k_0, \\ i_1 & j_1 & \ldots & k_1, \\ \ldots & \ldots & & \ldots, \\ i_r & j_r & \ldots & k_r. \end{matrix}$$

We shall assume that $i_0 < i_1 < \ldots < i_r$, $j_0 < j_1 < \ldots < j_r$, etc. If we can change the order of the factors in product in such a way that $i_\alpha \leq j_\alpha \leq \ldots \leq k_\alpha$ for $\alpha = 0, 1, \ldots, r$, we call it, after A. Young, a *standard product.*

LEMMA 1. *The set of standard products of degree* m *forms a* $\mathbf{K}$*-linear base of the module of homogeneous elements of degree* m *in* $\mathbf{K}[P]$.[4]

This follows readily from the quadratic relations between $P_{i_0 i_1 \ldots i_r}$ of the form

$$\sum_\sigma \operatorname{sgn}(\sigma) P_{i_0 \ldots i_{\alpha-1} i_\alpha^\sigma \ldots i_r^\sigma} P^\sigma_{j_0 \ldots j_\alpha \, j_{\alpha+1} \ldots j_r} = 0,$$

where σ runs over the $\binom{r+2}{\alpha+1}$ permutations of $(i_\alpha \ldots i_r j_0 \ldots j_\alpha)$ such that $i^\sigma_\alpha < \ldots < i^\sigma_r, j^\sigma_0 < \ldots < j^\sigma_\alpha$.

LEMMA 2. *Let*

$$F(P) = \sum c_{ij \ldots k} P_{i_0 i_1 \ldots i_r} P_{j_0 j_1 \ldots j_r} \ldots P_{k_0 k_1 \ldots k_r}$$

be a linear form in the standard products of degree m with coefficients in $\mathbf{K}$, *which vanishes after the substitution*

$$X_{ij} = X_{rr} = 0 \ (i > j).$$

Then $c_{ij \ldots k} = 0$ *for* $(i_0 i_1 \ldots i_r) \neq (01 \ldots r)$.[5]

This can be seen by examining the coefficient of the smallest power of T after the substitution $X_{ij} = Y_{ij}T^{(n+1-i)(n+1-i+j)}$. The same reasoning shows that the standard products of degree m are linearly independent over $\mathbf{K}$.

We are now in a position to prove our Theorem 1. Let G be a homogeneous element of the integral closure of $\mathbf{K}[P]$. Then G is necessarily contained in this ring when $\deg(G) \geq m$, where m does not depend on G.[6] We assume that m is the smallest integer of this nature, and we shall show that $m = 0$. Otherwise we can find a homogeneous element G of degree $m - 1$ in the integral closure of $\mathbf{K}[P]$, which is not contained in this ring. Since G is an element of $\mathbf{K}(P)$ which is integral over $\mathbf{K}[P]$, it is a fortiori an element of $\mathbf{K}(X)$ which is integral over $\mathbf{K}[X]$. Since $\mathbf{K}[X]$ is integrally closed, G can be written uniquely as a polynomial $G(X)$ in X_{ij} with coefficients in $\mathbf{K}$. On the other hand $P_{01 \ldots r}G$ is a homogeneous element of degree m in the integral closure of $\mathbf{K}[P]$, whence it can be written in the form

$$P_{01 \ldots r}G = \sum c_{ij \ldots k} I_{i_0 i_1 \ldots i_r}$$

as a linear combination of standard products of degree m, by Lemma 1. Let $\sum = P_{01 \ldots r}G'(P) + \sum''$ be the decomposition of $\sum$ such that $\sum''$ is the part of $\sum$ no term of which contains $P_{01 \ldots r}$. Then $P_{01 \ldots r}[G - G'(P)] = \sum''$ is a polynomial in X_{ij}, which vanishes after the substitution in Lemma 2. Therefore, by the same lemma, each coefficient $c_{ij \ldots k}$ of $\sum''$ must be zero; whence $G = G'(P)$, completing the proof.

THEOREM 2 (SEVERI'S THEOREM). *If W is a positive divisor on V, then W can be expressed as an intersection-product of V and a hypersurface in the ambient space of V.*

In fact, by the arithmetic normality of V, the intersection-products $V.H_m$ of V and hypersurfaces H_m of degree m in the ambient space of V not containing V form a complete linear system on V.[6] On the other hand, by a result of the author, the algebraic family of positive divisors of a given degree on V is linear,[7] whence our theorem follows.

2. *Invariant Theoretic Meaning.*—We shall now consider the relation of the arithmetic normality of the Grassmann variety and the first main theorem for the unimodular group. We prove the following theorem, which is stronger than Theorem 1 but in the proof of which we use Theorem 1:

THEOREM 3. *If $\mathbf{K}[X]$ is the ring of polynomials in X_{ij} with coefficients in* $\mathbf{K}$, *then* $\mathbf{K}[X] \cap \mathbf{K}(P) = \mathbf{K}[P]$.

We have only to show that every polynomial $F(X)$ contained in $\mathbf{K}(P)$ is integral over $\mathbf{K}[P]$. If $F(X) = \sum_\rho F_\rho(X)$ is the decomposition of $F(X)$ into homogeneous parts, then, for any quantity t, $F(tX) = \sum_\rho t^\rho F_\rho(X)$ is contained in $\mathbf{K}[X] \cap \mathbf{K}(P)$. Therefore, we may assume that $F(X)$ itself is a homogeneous polynomial in the $(r + 1)(n + 1)$ letters X_{ij}. Let V^* be the "representative cone" of V and let F^* be the function defined by $F(X)$ on V^*. If $F(X)$ is not integral over $\mathbf{K}[P]$, the function F^* has at least one polar variety W^*, which is also a cone, of dimension one

less than the dimension of V^*.[8] Let K be a field of definition of F^* containing the coefficients of $F(X)$; let $p'_{i_0 i_1 \ldots i_r}$ be the co-ordinates of a generic point p'^* of W^* over $\bar{K}$ and t a variable over $K(p'^*)$. Without loss of generality, we may assume that $p'_{01 \ldots r} \neq 0$. Then the set of $r+1$ points $x'_0, x'_1, \ldots, x'_r$ with homogeneous co-ordinates

$$x'_{\alpha j} = t p'_{01 \ldots \alpha-1, \alpha+1 \ldots r, j} \quad (0 \leq j \leq n)$$

for $\alpha = 0, 1, \ldots, r$ also determines a generic point p''^* of W^* over $\bar{K}$. However, $F(X)$ is finite at x', but F^* is not so at p''^*—a contradiction.

We note that Theorem 1 and Theorem 3 hold actually for any reference field. The following theorem is the *first main theorem of covariant vector invariants for* $SL(r+1)$:

THEOREM 4. *If $F(X)$ is a polynomial in X_{ij} with coefficients in a field K which is invariant under unimodular substitutions*

$$X'_{ij} = \sum_{k=0}^{r} a_{ik} X_{kj},$$

then $F(X)$ can be expressed as a polynomial in the $P_{i_0 i_1 \ldots i_l}$ with coefficients in K.

Proof: Since each homogeneous part of $F(X)$ is invariant, we may assume that $F(X)$ itself is homogeneous. In this case we conclude from the absolute irreducibility of a determinant that we have $F(\ldots, \sum_{k=0}^{r} a_{ik} X_{kj}, \ldots) = \det(a_{ik})^m F(X)$ with some nonnegative integer m. Therefore, $F(X)$ is an element of $K(P)$; whence our assertion follows by Theorem 4.

We note that Theorem 4 is apparently stronger than Theorem 3.

Finally, we note that $SL(r+1)$ is generated by the following set of $r(r+1)$ substitutions:

$$X'_i = X_i \qquad (i \neq i_0),$$
$$X'_{i_0} = X_{i_0} + t_{i_0 j_0} X_{j_0} \quad (i_0 \neq j_0).$$

In the case of characteristic zero the invariance of $F(X)$ under such substitutions is expressed simply by first "polarizations,"

$$\sum_{j=0}^{n} X_{j_0 j} \cdot \partial F / \partial X_{i_0 j} = 0 \quad (i_0 \neq j_0),$$

and this was the starting point of Severi's algebraic proof.

3. *Application to Chow Forms.*—We add here an application of the first main theorem for the unimodular group to a general theorem in algebraic geometry. Let A^r be a variety of r dimension in L^n, which is defined over a field K; let y be a generic point of A over K and let L^{n-1} be the hyperplane in the "dual space" of L^n, which corresponds to y. There exists one and only one subvariety T in the product of A and $(r+1)$-ply projective space $L^n \times L^n \times \ldots \times L^n$, which is defined over K, such that $T.(y \times L^n \times \ldots \times L^n) = y \times L^{n-1} \times \ldots \times L^{n-1}$. This variety does not depend on K and y when A is given. It is a simple matter to see that $pr_{L^n \times \ldots \times L^n}(T)$ is a variety of $(r+1)n - 1$ dimension. Therefore, if we denote by U_{ij} a set of letters to describe a variety in $L^n \times L^n \times \ldots \times L^n$, the above algebraic projection can be defined by a single absolutely irreducible $(r+1)$-ply homogeneous polynomial $F(U)$. This polynomial is uniquely determined up to a constant by A and is called the *Chow form of* A.[9] The Chow form of any positive cycle of r dimension in L^n can be defined by the corresponding product of Chow forms of its irreducible components. We shall now prove the following assertion:

THEOREM 5. *Any Chow form $F(U)$ of a positive cycle of dimension r and of degree m in L^n can be expressed as a homogeneous polynomial of degree m in the determinants*

$$\begin{vmatrix} U_{0i_0}U_{0i_1}\cdots U_{0i_r} \\ U_{1i_0}U_{1i_1}\cdots U_{1i_r} \\ \cdots\quad\cdots \\ U_{ri_0}U_{ri_1}\cdots U_{ri_r} \end{vmatrix}.$$

Proof: We may assume that the given cycle A^r is a variety. If $U'_{ij} = \sum_{k=0}^{n} a_{ik}U_{kj}$ is any unimodular substitution, $F'(U) = F(U')$ is also absolutely irreducible and defines the same variety A^r as $F(U)$. Therefore, $F(U)$ is at least a relative hence an absolute invariant, of $SL(r+1)$. On the other hand, it is a simple matter to see that $F(U)$ is of degree m in each $U_{\alpha j}$ $(0 \leq j \leq n)$ for $\alpha = 0, 1, \ldots, r$; whence our assertion follows by Theorem 4.

It follows from this theorem that we can define a one-to-one "birational" transformation from the set of positive divisors of dimension r and of degree m into the complete linear system of hypersurface sections of degree m of the Grassmann variety V.

* This work was supported by a research project at Harvard University, sponsored by the Office of Ordnance Research, U.S. Army.

[1] F. Severi, "Sulla varietà che rappresenta gli spazi subordinati di data dimensione immersi in uno spazo lineare," *Ann. di Mat.*, Vol. **24** (1915).

[2] H. Weyl, *Classical Groups* (Princeton: Princeton University Press, 1946), **1**, 45.

[3] A. Weil, *Foundations of Algebraic Geometry* ("Am. Math. Soc. Colloquium Publications," Vol. **29** [1946]).

[4] W. V. D. Hodge, "Some Enumerative Results in the Theory of Forms," *Proc. Cambridge Phil. Soc.*, **39**, 24–26 (1943).

[5] *Ibid.*, pp. 25–27.

[6] O. Zariski, "Complete Linear Systems on Normal Varieties and a Generalization of a Lemma of Enriques-Severi," *Ann. Math.*, **55**, 563 (1952).

[7] The theory of Grassmann varieties over fields of arbitrary characteristic will be published as a preliminary part of a theory of canonical classes of nonsingular projective models.

[8] Weil, *op. cit.*, p. 270.

[9] W. L. Chow and B. L. van der Waerden, "Ueber zugeordnete Formen und algebraische Systeme von algebraischen Mannigfaltigkeiten," *Math. Ann.*, Vol. **113** (1937).

---

*ON KÄHLER VARIETIES OF RESTRICTED TYPE**

BY K. KODAIRA

DEPARTMENT OF MATHEMATICS, PRINCETON UNIVERSITY

Communicated by S. Bochner, February 23, 1954

1. This is a preliminary report of a paper[1] concerning Kähler varieties of restricted type. A compact complex analytic variety V of complex dimension $n \geqq 2$ is called a *Kähler variety of restricted type*[2] or a *Hodge variety*[3] if V carries a Kähler metric $ds^2 = 2\sum g_{\alpha\bar{\beta}}\,(dz^\alpha\, d\bar{z}^\beta)$ such that the associated exterior form $\omega = i\sum g_{\alpha\bar{\beta}}\, dz^\alpha \wedge d\bar{z}^\beta$ belongs to the cohomology class of an *integral* 2-cocycle on V. In what follows, such a metric will be called a *Hodge metric*. It is well known that *every non-*

singular algebraic variety in a projective space carries a Hodge metric. Our main theorem asserts that the converse of this proposition holds. Namely, we have:

MAIN THEOREM. *A compact complex analytic variety is (bi-regularly equivalent to) a nonsingular algebraic variety imbedded in a projective space if it carries a Hodge metric.*

2. We give here a brief outline of the proof of our main theorem. Let V be a Hodge variety of complex dimension $n \geqq 2$ and let $\omega = i \sum g_{\alpha\bar{\beta}}\, dz^{\alpha}\, d\bar{z}^{\beta}$ be the closed 2-form associated with the Hodge metric on V. Then there exists a complex line bundle[4] F_1 over V whose characteristic class $c(F_1)$ contains ω. Letting $F = hF_1$ where h is a *sufficiently large positive integer*, we associate with F a "meromorphic" mapping Φ_F of V into a projective space $\mathfrak{S}$ in the following manner: Take a base $\{\varphi_0, \varphi_1, \ldots, \varphi_\lambda, \ldots, \varphi_d\}$ of the linear space $\Gamma(F)$ consisting of all holomorphic sections of the bundle F, and, letting $\{U_j\}$ be a finite covering of V, denote by $\varphi_{\lambda j}(z)$ the fiber co-ordinate of $\varphi_\lambda(z)$ over U_j, where z is a point in U_j. Then, considering $(\varphi_{0j}(z), \ldots, \varphi_{\lambda j}(z), \ldots, \varphi_{dj}(z))$ as the homogeneous co-ordinates of a point in the projective space $\mathfrak{S}$, we have

$$(\varphi_{0j}(z), \ldots, \varphi_{dj}(z)) = (\varphi_{0k}(z), \ldots, \varphi_{dk}(z))$$

for

$$z \in U_j \cap U_k.$$

Hence we can define a mapping Φ_F of V into $\mathfrak{S}$ by

$$z \to \Phi_F(z) = (\varphi_{0j}(z), \varphi_{1j}(z), \ldots, \varphi_{dj}(z)).$$

In view of a theorem of Chow,[5] every compact analytic subvariety of $\mathfrak{S}$ is an algebraic variety. Therefore, it is sufficient to prove that Φ_F is a bi-regular mapping of V into $\mathfrak{S}$.

In order to show that the mapping Φ_F is locally bi-regular at each point p on V, we construct the *quadratic transform*[6] $\tilde{V} = Q_p(V)$ of V with respect to the center p. The quadratic transformation Q_p maps p into a subvariety $S = Q_p(p)$ of $\tilde{V}$ which is (bi-regularly equivalent to) an $(n-1)$-*dimensional projective space*. Moreover, the complex line bundle $\{S\}$ over $\tilde{V}$ satisfies the relation[7]

$$\{S\}_S = -\{e\}, \tag{1}$$

where $\{S\}_S$ denotes the restriction of $\{S\}$ to S and $\{e\}$ is the complex line bundle over S determined by the hyperplane e on S. We note that the mapping Q_p is bi-regular between $V - p$ and $\tilde{V} - S$. Now, let $\tilde{F} = Q_p F$ be the complex line bundle over $\tilde{V}$ induced by F in a canonical manner. Since the bundle $\tilde{F}$ is trivial over S, Q_p induces *the isomorphism*

$$\Gamma(F) \cong \Gamma(\tilde{F}). \tag{2}$$

Denoting by K the canonical bundle over V, we infer that the canonical bundle $K(\tilde{V})$ over $\tilde{V}$ is given by

$$K(\tilde{V}) = \tilde{K} + (n-1)\{S\}, \tag{3}$$

where $\tilde{K} = Q_p K$. Letting

$$E_m = \tilde{F} - m\{S\}, \qquad m = 0, 1, 2,$$

we therefore have

$$E_m - K(\tilde{V}) = \tilde{F} - \tilde{K} - (n + m - 1)\{S\}. \tag{4}$$

By means of (1) it can be shown that the characteristic class $c(-\{S\})$ of $-\{S\}$ contains a closed real (1, 1)-form σ whose restriction σ_S to S is equal to the fundamental form on the projective space S. With the help of this form, σ, we infer from (4) that $c(E_m - K(\tilde{V}))$ contains a closed *positive* (1, 1)-form for $m = 1, 2$. Consequently, we have[8]

$$H^1(\tilde{V}, \Omega(E_m)) = 0, \qquad \text{for } m = 1,2, \tag{5}$$

where $\Omega(E_m)$ denotes the sheaf over $\tilde{V}$ of germs of holomorphic sections of E_m. Since the restriction $\tilde{F}_S$ is a trivial bundle over S, we infer from (1) that $(E_m)_S = \{me\}$. Hence we have the exact sequences,

$$0 \to \Omega(E_{m+1}) \to \Omega(E_m) \xrightarrow{r} \Omega(\{me\}) \to 0,$$

where r denotes the restriction map to S. Combined with (5), these sequences yield

$$0 \to \Gamma(E_1) \to \Gamma(\tilde{F}) \xrightarrow{r^*} C_S \to 0,$$

$$0 \to \Gamma(E_2) \to \Gamma(E_1) \xrightarrow{r^*} \Gamma(\{e\}) \to 0, \tag{6}$$

where C_S is the space of constants on S. Now it is easy to derive from (2) and (6) that the mapping Φ_F is bi-regular in a neighborhood U_p of p, while p is an arbitrary point on V. Consequently, Φ_F *is a locally bi-regular mapping of* V *into* $\mathfrak{S}$.

To prove that $\Phi_F(p) \neq \Phi_F(q)$ for any pair of points p,q, $p \neq q$, on V, we form the quadratic transform $\tilde{V} = Q_pQ_q(V)$ and let

$$E' = \tilde{F} - \{S\},$$

$$E'' = \tilde{F} - \{S\} - \{T\},$$

where $\tilde{F} = Q_pQ_qF$, $S = Q_pQ_q(p)$, and $T = Q_pQ_q(q)$. Then we get, as above, the exact sequences

$$0 \to \Gamma(E') \to \Gamma(\tilde{F}) \xrightarrow{r^*_S} C_S \to 0,$$

$$0 \to \Gamma(E'') \to \Gamma(E') \xrightarrow{r^*_T} C_T \to 0, \tag{7}$$

where r^*_S and r^*_T are the restriction maps to S and T, respectively. On the other hand, we have the isomorphism $\Gamma(F) \cong \Gamma(\tilde{F})$. From these results it follows that $\Phi_F(p) \neq \Phi_F(q)$. Thus we conclude that $z \to \Phi_F(z)$ *is a one-to-one bi-regular mapping of* V *into* $\mathfrak{S}$. This completes the proof of our main theorem.

3. Now we mention several applications of our main theorem. First, let V be a compact complex analytic manifold with a Hermitian metric $\sum g_{\alpha\bar{\beta}}(dz^\alpha\, d\bar{z}^\beta)$. Letting

$$R_{\alpha\bar{\beta}} = -\frac{\partial^2 \log g}{\partial z^\alpha\, \partial \bar{z}^\beta}, \text{ where } g = \det(g_{\alpha\bar{\beta}}),$$

be the Ricci curvature, we infer readily that the metric

$$ds^2 = -[\text{or} +] \sum \left(\frac{R_{\alpha\bar{\beta}}}{\pi}\right) (dz^\alpha\, d\bar{z}^\beta)$$

is a Hodge metric if it is positive definite. Hence we obtain the following theorem:

THEOREM. *Every compact Hermitian manifold with negative (or positive) definite Ricci curvature is a nonsingular algebraic variety imbedded in a projective space.*

4. Second, let $\mathfrak{B}$ be a bounded domain on the space of n complex variables and let Δ be a discontinuous group of analytic automorphisms of $\mathfrak{B}$ such that $\mathfrak{B}$ is compact with respect to Δ. Moreover, assume that no element of Δ except the identity has a fixed point in $\mathfrak{B}$. Then the factor space $V = \mathfrak{B}/\Delta$ is a compact analytic manifold. Now it can be shown that the Bergmann metric,

$$ds^2 = \frac{1}{\pi} \sum \frac{\partial^2 \log K(z,\bar{z})}{\partial z^\alpha \partial \bar{z}^\beta} (dz^\alpha\, d\bar{z}^\beta),$$

of $\mathfrak{B}$ induces on $V = \mathfrak{B}/\Delta$ a Hodge metric, where $K(z,\bar{z})$ denotes Bergmann's kernel function for the domain $\mathfrak{B}$. Consequently, we obtain the following theorem:

THEOREM.[9] *The compact analytic manifold $V = \mathfrak{B}/\Delta$ is a nonsingular algebraic variety imbedded in a projective space.*

This theorem implies the existence of "sufficiently many" automorphic functions on $\mathfrak{B}$ with respect to Δ.

5. Finally, we consider projective bundles over algebraic varieties. Let B be an analytic fibre bundle over a Hodge variety V whose fibre is a complex projective space and whose structure group is the group of projective transformations. Then it can be shown that B is also a Hodge variety. Hence we obtain the following theorem:

THEOREM. *Every projective bundle over a nonsingular algebraic variety in a projective space is also a nonsingular algebraic variety imbedded in a projective space.*[10]

* This work was supported by a research project at Princeton University sponsored by the Office of Ordnance Research, U.S. Army.

[1] K. Kodaira, "On Kähler Varieties of Restricted Type," to appear in the *Ann. Math.*

[2] W. V. D. Hodge, "A Special Type of Kähler Manifolds," *Proc. London Math. Soc.*, **1**, 104–117 (1951).

[3] A. Weil, "On Picard Varieties," *Am. J. Math.*, **74**, 865–894 (1952).

[4] K. Kodaira and D. C. Spencer, "Groups of Complex Line Bundles over Compact Kähler Varieties," these PROCEEDINGS, **39**, 868–872 (1953).

[5] W. L. Chow, "On Compact Complex Analytic Varieties," *Am. J. Math.*, **71**, 893–914 (1949).

[6] See F. Hirzebruch, "Über vierdimensionale Riemannsche Flächen mehrdeutiger analytischer Funktionen von zwei komplexen Veränderlichen," *Math. Ann.*, **126**, 1–22 (1953).

[7] This relation is due to F. Hirzebruch.

[8] K. Kodaira, "On a Differential-geometric Method in the Theory of Analytic Stacks," these PROCEEDINGS, **39**, 1268–1273, Theorem 3 (1953).

[9] A. Borel has communicated to the author the fact that a similar result has been obtained recently by J.-P. Serre.

[10] The author wishes to express his sincere thanks to A. Borel for valuable suggestions concerning these applications.

*ON THE DERIVATION OF THE EQUATIONS OF HYDRODYNAMICS FROM STATISTICAL MECHANICS**

By Charles B. Morrey, Jr.

UNIVERSITY OF CALIFORNIA, BERKELEY

Communicated by G. C. Evans, March 23, 1954

1. *Introduction.*—The object of this paper is to use several methods to derive the laws of the mass motion of a "continuous medium" from the study of the motions of the finite systems of particles described below. A long paper containing proofs of most of the results presented here and a further discussion of them has been submitted to *Communications on Pure and Applied Mathematics.*

Three methods are used to pass from the discrete to the continuous: (1) a study of the Fourier-Stieltjes transforms of the finite distributions of mass, momentum, and energy for the particle distributions themselves; (2) the Gibbs approach, in which a continuous family of particle distributions is used; and (3) the reduction of a problem of finding time averages to that of finding space averages, using the Ergodic theorem. In all cases, a limiting process is carried through, in which the number, N, of particles is allowed to become infinite.

The first method is used to establish the existence of certain limiting distributions (as $N \to \infty$) which vary continuously with time. The Gibbs approach leads to the usual system of integrodifferential equations connecting the distributions of co-ordinates and velocities for clusters of particles.[1] A change of variables is made which enables us to allow $N \to \infty$ in these equations. The resulting infinite equilibrium system is solved exactly *without recourse to the usual superposition assumption* assuming the usual form of the equilibrium solutions. The usual hydrodynamic equations, without viscosity and heat conduction, follow, together with expressions for the pressure, energy, and density in terms of the potential function, the temperature, and another parameter. This form allows for the possible inclusion of both the gaseous and liquid states, but the writer does not know as yet whether both states are actually included. The third method justifies the assumed form of the equilibrium solutions and also leads to the entropy in a natural way.

2. *Particle Distributions.*—The finite systems of particles considered consist of N identical point-mass particles in Euclidean space, each of mass m_N, any two of which repel one another with a force $-m_N^2\,\Phi_N'(r)$ (if $\Phi_N'(r) > 0$, the particles attract one another). At each time t, we define the distributions

$$D_N^1(t; R) = m_N \nu_N(t; R) \qquad \text{(mass)};$$

$$D_N^{1+\alpha}(t; R) = m_N \sum_{x_j(t)\in R} u_j^\alpha(t)\,, \qquad \alpha = 1, 2, 3 \qquad \text{(momentum)};$$

$$D_N^5(t; R) = {}^1\!/_2 m_N \sum_{x_j(t)\in R} \left[|u_j(t)|^2 + m_N \sum_{k=1}^{N}{}' \Phi_N(|x_j - x_k|) \right] \qquad \text{(energy)}. \qquad (1)$$

In equation (1), $\nu_N(t; R)$ is the number of particles in the set R at the time t; $D_N^1(t; R)$ is the mass of those particles; $D_N^{1+\alpha}(t; R)$ are the components of the total momentum of those particles, and $D_N^5(t; R)$ is their total energy; the prime in the inner summation in D_N^5 requires that $k \neq j$. The distributions $P_{N1}(t; R) = M_N \Pi_{N1}(t; R)$ of co-ordinates and velocities are defined similarly, R denoting a set in 6-space; here

$\Pi_{N1}(t; R) = N^{-1}\nu_N(t; R)$ is the probability distribution. The higher-order distributions $P_{Nl}(t; S)$ and $\Pi_{Nl}(t; S)$ are defined for sets S in the 6-space with

$$P_{Nl}(t; S) = M_N{}^l N! \Pi_{Nl}(t; S)/N^l(N - l)!, \quad M_N = Nm_N. \tag{2}$$

3. *The Normalized Potential Function.*—As was pointed out above, we wish to allow $N \to \infty$ "normalizing" m_N and $\Phi_N(r)$, so as to retain the essential features of the system. This is done by choosing a scale length $\sigma_N \to 0$ as $N \to \infty$ and defining

$$m_N = D\sigma_N{}^3, \quad m_N\Phi_N(r) = K\Phi(r/\sigma_N), \tag{3}$$

in which D, K, and Φ are independent of N.

The potential Φ is allowed to be negative for some r but must satisfy

$$\text{(i)} \lim_{r \to 0+} r^\alpha \Phi(r) = +\infty, \lim_{r \to +\infty} r^\alpha \Phi(r) = 0 \text{ for some } \alpha > 3;$$

(ii) $\Phi'(r)$ is continuous for $r > 0$ with

$$\limsup_{r \to 0+} |r\,\Phi'(r)|/\Phi(r) = \delta < +\infty; \lim_{r \to +\infty} r^{\alpha+1}\Phi'(r) = 0. \tag{4}$$

The conditions (4) obviously imply that $\Phi(r) \geqq 0$ or else has a negative minimum $-m$. It is also true, but is considerably less obvious, that those conditions imply the existence of numbers L and $L_1 \geqq 0$, such that

$$N^{-1} \sum_{j,k=1}{}' \alpha_{jk}\Phi(|x_j - x_k|) \geqq -L, \quad (\alpha_{jk} = \alpha_{kj}, \alpha_{jk} = 0 \text{ or } 1), \tag{5}$$

$$N^{-1} \sum_{j,k=1}{}' |x_j - x_k| \cdot |\Phi'(|x_j - x_k|)| \leqq (\delta + 1)N^{-1} \sum_{j,k=1}{}' \Phi(|x_j - x_k|) + L_1, \tag{6}$$

independently of N.

4. *Some Limit Theorems.*—The total mass, total energy, and second x moment of P_{N1} are given by

$$M_N = Nm_N, \; E_N = {}^1/_2 m_N \sum_{j=1}^{N} \left[|u_j|^2 + K \sum_{k=1}{}' \Phi(|x_j - x_k|/\sigma_N) \right],$$

$$C_N = m_N \sum_{j=1}^{N} |x_j|^2, \tag{7}$$

respectively. For any given system, M_N and E_N are independent of time. From equation (5), it follows that the total kinetic and potential energies satisfy

$${}^1/_2 m_N \sum_{j=1}^{N} |u_j|^2 \leqq E_N + {}^1/_2 M_N KL, \quad {}^1/_2 K m_N \sum_{j,k=1}{}' \Phi(|x_j - x_k|/\sigma_N) \leqq E_N \tag{8}$$

for all times. From this, it is easy to see that

$$C_N(t) \leqq \{ [C_N(t_0)]^{1/2} + (2E_N + M_N KL)^{1/2} |t - t_0| \}^2 \tag{9}$$

for all t and t_0.

We let $\psi_N{}^\gamma(t; y)$ be the Fourier-Stieltjes transform of $D_N{}^\gamma$ ($\gamma = 1, \ldots, 5$); then

$$\psi_N{}^1(t;\, y) = m_N \sum_{j=1}^{N} \exp\,[iy\cdot x_j(t)\,],$$

$$\psi_N{}^{1+\alpha}(t;\, y) = m_N \sum_{j=1}^{N} u_j{}^{\alpha}(t)\cdot\exp[iy\cdot x_j(t)\,], \qquad \alpha = 1, 2, 3, \quad (10)$$

for instance. Suppose we consider a sequence of particle distributions in which $N \to \infty$, but M_N, E_N, and C_N remain bounded independently of N at some instant t_0. Then we see that $\psi_N{}^1, \ldots, \psi_N{}^4$ and their first derivatives with respect to t and the y^{α} are continuous and are uniformly bounded on any bounded part of the $(t; y)$ space independently of N, and also

$$\psi_{Nt}{}^1(t;\, y) = iy^{\alpha}\psi_N{}^{1+\alpha}(t;\, y) \qquad (\alpha \text{ summed}), \quad (11)$$

which is formally the Fourier transform of the equation of continuity. Accordingly, by Ascoli's theorem, *we may extract a subsequence of the N such that the functions $\psi_N{}^{\gamma}$ converge uniformly on any bounded part of $(t; y)$ space to limit functions $\psi^{\gamma}(\gamma = 1, \ldots, 4)$.*

As to the $\psi_N{}^5$, we see only that they are uniformly bounded. However, the derivatives $\psi_{Nt}{}^5$ do not have $\sigma_N{}^{-1}$ as a factor, although they involve third moments and cross-moments not proved to be bounded in time. The equations of motion of the Nth system become, using equation (3),

$$\dot{x}_j{}^{\alpha}(t) = u_j{}^{\alpha}(t), \quad \dot{u}_j{}^{\alpha} = -K\sigma_N{}^{-1} \sum_{k=1}{}' \Phi'(|x_j - x_k|/\sigma_N)\cdot(x_j{}^{\alpha} - x_k{}^{\alpha})/|x_j - x_k|, \quad (12)$$

so that further time derivatives of $\psi_N{}^{1+\alpha}$ have $\sigma_N{}^{-1}$ as a factor.

5. *The Gibbs Approach.*—The limiting theorem stated above focuses attention on the determination of the limiting distributions, which will be some sort of quasi-stable distributions, since they vary continuously with the time. In order to determine them, we begin by employing the device of Gibbs for passing from the discrete to the continuous. All the distributions P_{Nl}, Π_{Nl}, $D_N{}^{\gamma}$, etc., are thereby replaced by distributions with density functions p_{Nl}, π_{Nl}, etc., with whatever differentiability and integrability properties are desired. All these functions are symmetric in the indices occurring in them, and π_{NN} (and, of course, p_{NN}) satisfies Liouville's equation in the phase space. The π_{Nl} satisfy

$$\pi_{Nl}(t;\, x_1, u_1; \ldots ; x_l, u_l) = \int_{-\infty}^{\infty} \pi_{N,\, l+1}(t;\, x_1, u_1; \ldots ; x_{l+1}, u_{l+1})\,dx_{l+1}du_{l+1} \quad (13)$$

and hence satisfy a system of equations like that in Born and Green, *op. cit.*, pages 4–6, obtained from Liouville's equation using equation (13).

In order to carry out a limiting process as $N \to \infty$, it is convenient to use the p_{Nl} rather than the π_{Nl} (see eq. [2] for the connection) and to introduce new variables $(x, \xi_1, \ldots, \xi_l, \ldots ; u, v_1, \ldots, v_l, \ldots)$ and new functions ϕ_{Nl} by

$$\varphi_{Nl}(t;\, x, u;\, \xi_1, v_1; \ldots ; \xi_{l-1}, v_{l-1}) =$$
$$p_{Nl}(t;\, x, u;\, x + \sigma_N\xi_1, u + v_1; \ldots ; x + \sigma_N\xi_{l-1}, u + v_{l-1}). \quad (14)$$

The φ_{Nl} satisfy the following system of equations:

$$\sigma_N(\varphi_{N\,l\,u} + u^\alpha \varphi_{Nx^\alpha}) + B_{l0}{}^\alpha(\xi)\varphi_{N\,l0^\alpha} + \sum_{j=1}^{l-1} [v_l{}^\alpha \varphi_{Nl\xi_j{}^\alpha} + B_{lj}{}^\alpha(\xi)\varphi_{N\,lv_j{}^\alpha}] =$$

$$-F^\alpha_{Nlou^\alpha} - \sum_{j=1}^{l-1} (F^\alpha_{Nljv_j{}^\alpha} - F^\alpha_{Nlov_j{}^\alpha}), \quad l = 1, \ldots, N \quad (\alpha \text{ summed}):$$

$$F_{N\,lj}{}^\alpha = D^{-1}K \int_{-\infty}^{\infty} [\Phi'(|\xi_j - \xi_l|)\cdot(\xi_l{}^\alpha - \xi_j{}^\alpha)\varphi_{N,\,l+1}/|\xi_j - \xi_l|]d\xi_l dv_l, \quad j = 0, \ldots, l-1;$$

$$B_{l0}{}^\alpha(\xi) = K \sum_{j=1}^{l-1} \Phi'(|\xi_j|)\xi_j{}^\alpha/|\xi_j|, \quad B_{lj}{}^\alpha = -H_{l\xi_j{}^\alpha} - B_{l0}{}^\alpha, \qquad j = 1, \ldots, l-1;$$

$$H_l = \frac{K}{2} \sum_{j,\,k=0}^{l-1}{}' \Phi(|\xi_j - \xi_k|), \quad \xi_0 = 0, \tag{15}$$

where the prime in the last sum requires $k \neq j$ (as usual); if $l = N$, the right side of the first equation is zero. Except for the factor σ_N in the first term, the coefficients in equation (15) are *independent of* N.

The distributions (1) now have the form

$$D_N{}^1(t; R) = \int_R \rho_N(t; x)dx, \quad D_N{}^{1+\alpha}(t; R) = \int_R \rho_N(t; x)\tilde{u}_N{}^\alpha(t; x)dx,$$

$$D_N{}^5(t; R) = \int_R e_N(t; x)dx, \quad e_N = \rho_N[|\tilde{u}_N|^2/2 + \epsilon_N(t; x)], \tag{16}$$

where ϵ_N is the internal energy per unit mass and $\tilde{u}_N$ is the mass-velocity vector. In terms of the φ_{Nl}, we obtain

$$\rho_N(t; x) = \int_{-\infty}^{\infty} \varphi_{N1}(t; x, u)du, \quad \rho_N(t; x)\tilde{u}_N{}^\alpha(t; x) = \int_{-\infty}^{\infty} u^\alpha \varphi_{N1}(t; x, u)du,$$

$$e_N(t; x) = {}^1/_2 \int_{-\infty}^{\infty} |u|^2 \varphi_{N1}\, du + {}^1/_2 D^{-1}K \int_{-\infty}^{\infty} \Phi(|\xi_1|)\varphi_{N2}(t; x, u; \xi_1, v_1])\, du\, d\xi_1\, dv_1. \tag{17}$$

Using equations (15) and the symmetry of the p_N in the indices, equations are found for the time derivatives of ρ_N, $\rho_N\tilde{u}_N{}^\alpha$, and e_N in terms of other derivatives of these functions and other functions defined by integrals something like those in equation (17), but involving σ_N and $\sigma_N{}^{-1}$ in such a way that they tend to limits as $N \to \infty$ and $\sigma_N \to 0$, provided that φ_{N1} and φ_{N2} and their derivatives converge properly to their limits.

6. *Determination of the Limiting Functions φ_l in the Gaseous (and Liquid?) States.*—Since the limiting distributions vary reasonably with time and space, we allow $N \to \infty$ and $\sigma_N \to 0$ in equation (15). There results an infinite system in which t and x^α do not enter explicitly. The number of solutions of the resulting system is obviously infinite. So, on the basis of "experience," we look for solutions φ_l of the form

$$\varphi_l = (A/\pi)^{3l/2} \exp\left\{-A\left[\sum_{j=0}^{l-1} |u + v_j|^2 + K \sum_{j,\,k=0}^{l-1}{}' \Phi(|\xi_j - \xi_k|)\right]\right\} \times$$

$$\rho^l\, g_l(\xi_1, \ldots, \xi_{l-1};\ \rho, A), \quad \rho = \rho(t; x), \quad A = A(t; x) = 1/2kT \quad (v_0 = 0) \tag{18}$$

(see Born and Green, *op. cit.*, p. 6, for instance), in which we require that $g_l \to 1$ as all the $|\xi_j|$ and $|\xi_j - \xi_k| \to \infty$; we make this last requirement since it is easy to see from the definitions that, for particle systems,

$$P_{N1}(t;\, R_1 \times \ldots \times R_l) = \prod_{j=1}^{l} P_{N1}(t;\, R_j), \tag{19}$$

provided that the R_j are disjoint ($R_1 \times \ldots \times R_l$ is all the $[x_1, u_1; \ldots; x_l, u_l]$ in which $[x_j, u_j] \in R_j$).

Equations (18) require that the g_l satisfy

$$g_{lx_j^\alpha} = -D^{-1}\rho \int_{-\infty}^{\infty} g_{l+1} \cdot \frac{\partial}{\partial \xi_{j\alpha}} [1 - \exp(-2A\,\Omega_l)]d\xi_l,\ \Omega_l = K \sum_{j=0}^{l-1} \Phi(|\xi_l - \xi_j|),$$

$$\int_{-\infty}^{\infty} g_{l+1} \cdot \frac{\partial}{\partial \xi_{l\alpha}} [1 - \exp(-2A\,\Omega_l)]d\xi_l = 0, \quad \alpha = 1, 2, 3; l = 1, 2, \ldots (\xi_0 = 0); \tag{20}$$

Without the use of the superposition approximation, it is found that equations (20) have the solutions

$$g_l(\xi_1, \ldots, \xi_{l-1};\, \rho,\, A) = [E(\rho,\, A)]^l \sum_{q=0}^{\infty} g_{lq}(\xi_0, \ldots, \xi_{l-1};\, A) \cdot (-y)^q/q!\,(\xi_0 = 0), \tag{21}$$

where y and E are possibly multiple-valued functions of ρ and A, defined by

$$\rho/D = yG_1(y;\, A), \quad E = [G_1(y;\, A)]^{-1}, \quad G_1(y,\, A) = \sum_{p=0}^{\infty} g_{1p}(A) \cdot (-y)^p/p!, \tag{22}$$

and the $g_{lp}(\zeta_1, \ldots, \zeta_l;\, A)$ are functions defined by the recursion relations

$$g_{lp}(\zeta_1, \ldots, \zeta_l;\, A) = \sum_{q=1}^{p} \frac{p!}{q!(p-q)!} \int_{-\infty}^{\infty} k_{lq}(\zeta;\, \eta_1, \ldots, \eta_q;\, A) g_{q,\, p-q}(\eta_1, \ldots, \eta_q;\, A)d\eta, \tag{23}$$

where

$$k_{lq}(\zeta;\, \eta_1, \ldots, \eta_q;\, A) = K_q(\eta_1, \ldots, \eta_q;\, A) \cdot \prod_{r=1}^{q} k_{l1}(\xi;\, \eta_r;\, A),$$

$$K_q(\eta_1, \ldots, \eta_q;\, A) = \exp\left[-KA \sum_{j,\, k=1}^{q}{}' \Phi(|\eta_j - \eta_k|)\right],$$

$$k_{l1}(\zeta_1, \ldots, \zeta_l;\, \omega;\, A) = 1 - \exp\left[-2KA \sum_{j=1}^{l} \Phi(|\zeta_j - \omega|)\right]; \tag{24}$$

if $l = 1$, the $g_{lp}(\zeta_1;\, A)$ are just constants $g_{1p}(A)$, as indicated in equation (22). The series in equations (21) and (22) converge, at least if

$$|y| \leqq [eP(A)\tau(A)]^{-1}, \quad \tau(A) = \tau_0(A) + \tau_1(A),$$
$$P(A) = \exp(2KAL), \quad \tau_i = \int_{-\infty}^{\infty} \{1 - \exp[-2KA\Phi_i(|\eta|)]\}d\eta, \quad i = 1, 2,$$
$$\Phi(r) = \Phi_0(r) - \Phi_1(r), \quad \Phi_0(r) = {}^1/_2[|\Phi(r)| + \Phi(r)] \geqq 0. \tag{25}$$

For these values of y, the solutions given by equation (21) are unique. From equation (22), we see that, for some values of A, two or more different values of y may yield the same values of ρ, so that E and the g_l may be multiple-valued.

7. *Justification of the Form (18): Conclusion.*—For this purpose, we begin by reverting to the particle distributions and consider a sequence of such at a given instant which converge to certain smooth distributions. For each N, we "include practically all" of the mass, momentum, and energy of the Nth distribution in a finite number of disjoint cells $R_1, \ldots, R_{Q_N}$ where $Q_N \to \infty$ rather slowly. We

compute, for that instant, the distributions P_{Nl}, by averaging the P_{Nl} for all particle distributions whose five distributions $D_N{}^{\gamma}(t_0; R)$ coincide with the given one for each R_j, $j = 1, \ldots, Q_N$. This problem is simplified to that of computing the equilibrium distribution of ν particles in a cell R with given values of ρ and ϵ. This reduces to the problem of computing ratios of the measures of sets on certain manifolds. This is a rather standard type of problem but is too long to describe here. The limiting process is carried through, allowing $\nu \to \infty$, and the results are seen to be of the form (18), in which the g_l satisfy equation (20).

The ordinary hydrodynamic equations, without viscosity or heat conduction, follow easily in the usual way from the limiting equations in section 5, above; expressions are obtained for the pressure p and for ϵ in terms of Φ, ρ, and A. The analysis of the procedure in the immediately preceding paragraph yields further information about the equation of state "in the large" and introduces the entropy function in a natural way.

* The results presented in this paper were obtained by the writer under a contract with the Office of Naval Research.

[1] M. Born and H. S. Green, *A General Kinetic Theory of Liquids* (Cambridge: At the University Press, 1949), p. 5.

DETERMINATION OF THE PLASTIC YIELD CONDITION AS A VARIATIONAL PROBLEM

By T. Y. THOMAS

NAVAL RESEARCH LABORATORY, WASHINGTON, D.C.

Communicated February 11, 1954

1. *Introduction.*—Dr. G. R. Irwin has suggested that plastic flow occurs in such a way as to minimize a certain energy integral taken over the region of flow. This appears to us to be a particularly significant idea, and we have accordingly attempted to derive certain of its consequences in the following communication. In so doing, we have based our work on the *deviation tensors* of stress and strain rate, in view of the generally recognized fact that plastic phenomena are more or less independent of the effects of hydrostatic pressure. Thus, if u_α denote the components of velocity and $\sigma_{\alpha\beta}$, the ordinary stress components, we have[1]

$$e_{\alpha\beta} = \frac{1}{2}(u_{\alpha,\beta} + u_{\beta,\alpha}); \qquad \eta_{\alpha\beta} = e_{\alpha\beta} - \frac{e_{ii}}{3}\delta_{\alpha\beta}; \qquad s_{\alpha\beta} = \sigma_{\alpha\beta} - \frac{\sigma_{ii}}{3}\delta_{\alpha\beta} \quad (1)$$

for the components of the strain-rate tensor e, its deviation η, and the stress-deviation tensor s. Actually, the quantities u_α can also be interpreted as displacements in the following discussion, in which case the tensor e is the ordinary strain tensor.

Now in the elastic domain the quantity $s_{\alpha\beta}s_{\alpha\beta}$ represents, apart from a constant factor depending on the elastic moduli of the material, the distortion energy per unit volume, i.e., the energy per unit volume stored in the body as the result of change in shape as distinct from change in volume. It would be natural to consider the distortion energy in the plastic regime and to deduce the effect of minimiz-

ing the integral of this energy as one of the characteristic conditions governing the plastic flow. However, it is obvious that the expression $s_{\alpha\beta}s_{\alpha\beta}$, derived from purely elastic theory, may not represent, to within a constant factor of proportionality, the plastic distortion energy, except perhaps as a first approximation, since, in the plastic, as distinct from the elastic, case, we are concerned with an irreversible process involving the phenomenon of permanent set, as well as other factors which it may be necessary to take into account in any consideration of energy content.

In the following article we have considered the effect on the plastic flow of the requirement that the integral of an arbitrary differentiable (invariant) function $\psi(s, \eta)$ of the components of the deviation tensors s and η have a stationary value over the plastic domain. Strong differential conditions for the stationary character of this integral are derived in section 2 of this article by the method of virtual displacements. Now, when these conditions are combined with the usual dynamical conditions of equilibrium and the usual relation expressing the incompressibility of the medium, we have an *overdetermined system*. In section 4 we have shown how this overdetermination can be removed, and, in so doing, we are led to an extended system of differential equations for the function ψ; this system remains valid for the case of compressible material (sec. 3). Treating this extended system on the basis of a consideration involving canonical co-ordinates, it is shown in section 6 that ψ can, quite generally, be a function at most of the trace η_{ii} of the tensor η. But $\eta_{ii} \equiv 0$, from the definition of the deviation tensor η, and hence we arrive at the conclusion that $\psi = \text{Const.}$ *In other words, ψ must reduce to a constant as a function of the components of the η tensor.* This result is of the nature of a yield condition, as emphasized in section 7.

If, in particular, we assume that the tensors s and η are proportional and if we take the function ψ to be given by the expression $s_{\alpha\beta}s_{\alpha\beta}$, we arrive at the quadratic yield condition and stress-strain relations used by von Mises. This same yield condition is also obtained by the selection $\psi = s_{\alpha\beta}s_{\alpha\beta}$ and the most general relation between the tensors s and η, considered in section 8.

According to the above result, the volume integral of the function ψ must remain unchanged by virtual displacements. This situation, which differs from that in the classical theory of elasticity, where the integral of potential energy attains an absolute minimum in a strong sense, must be viewed as a characteristic property of the plastic flow. That this result is not trivial in spite of its simplicity is evident from the analysis needed for its demonstration, as well as from the fact that it depends essentially on the imposed requirement of nonoverdetermination of the differential relations and on the assumption that the stress deviation s is an invariant of the deviation tensor η. Now this latter condition is not met in the type of relation assumed between the tensors s and η of certain theories of plasticity, e.g., the Prandtl-Reuss theory. *Since the stress-strain relationship is closely connected with the nature of the function ψ to be used in this variational procedure, an investigation, based on physical considerations, of some of the most likely choices of the function ψ may throw some light on these matters.*

2. *Differential Conditions.*—Consider a medium under the action of surface forces, and denote by u_α the components of the resulting displacement or velocity of the material particles. If the surface forces act only on a part Σ_1 of the boundary, it is assumed that over the remaining part Σ_2 of the boundary the values of the u_α

are assigned. Now consider a change δu_α in the quantities u_α, with δu_α arbitrary, subject to conditions of continuity and differentiability and to the above boundary conditions, i.e., $\delta u_\alpha = 0$ on the part Σ_2 of the boundary where the u_α are assigned. Such a displacement δu is called a *virtual displacement*.

For simplicity we suppose the medium under discussion to be entirely in the plastic regime.[2] We now impose the condition that

$$\delta \int_V \psi(s, \eta) dV = 0 \tag{2}$$

under the virtual displacement δu, where the integration is over the entire volume V of the plastic body. In treating condition (2) above, it will be assumed that s is a tensor invariant of the tensor η, i.e., $s_{\alpha\beta} = s_{\alpha\beta}(\eta)$. The exact form of this relationship will be considered later, as it will not be needed in the immediate discussion. Now we have

$$\delta\eta_{\sigma\rho} = \frac{1}{2}(\delta u_{\sigma,\rho} + \delta u_{\rho,\sigma}) - \frac{1}{3}\delta u_{t,t}\delta_{\sigma\rho} = \frac{1}{2}\left(\frac{\partial \delta u_\sigma}{\partial x^\rho} + \frac{\partial \delta u_\rho}{\partial x^\sigma}\right) - \frac{1}{3}\frac{\partial \delta u_t}{\partial x^t}\delta_{\sigma\rho},$$

where the differentiation is with respect to rectangular co-ordinates x. Hence it is easily seen that condition (2) above can be written in the form

$$\int_V \frac{\partial \psi}{\partial \eta_{\sigma\rho}} \frac{\partial \delta u_\sigma}{\partial x^\rho} dV - \frac{1}{3}\int_V \frac{\partial \psi}{\partial \eta_{\sigma\sigma}} \frac{\partial \delta u_\rho}{\partial x^\rho} dV = 0, \tag{3}$$

where the function ψ in this equation is that which results from the original function $\psi(s, \eta)$ by the above-mentioned substitution $s = s(\eta)$. But equation (3) can also be written

$$\left.\begin{aligned} &\int_V \left(\frac{\partial \psi}{\partial \eta_{\sigma\rho}} \delta u_\sigma\right)_\rho dV - \int_V \left(\frac{\partial \psi}{\partial \eta_{\sigma\rho}}\right)_\rho \delta u_\sigma \, dV \\ &- \frac{1}{3}\int_V \left(\frac{\partial \psi}{\partial \eta_{\sigma\sigma}} \delta u_\rho\right)_\rho dV + \frac{1}{3}\int_V \left(\frac{\partial \psi}{\partial \eta_{\sigma\sigma}}\right)_\rho \delta u_\rho \, dV = 0 \end{aligned}\right\}. \tag{4}$$

Now the first and third integrals in this equation can be expressed as surface integrals by the application of Green's theorem. We thus obtain

$$\int_V \left[\left(\frac{\partial \psi}{\partial \eta_{\sigma\rho}}\right)_\rho - \frac{1}{3}\left(\frac{\partial \psi}{\partial \eta_{\rho\rho}}\right)_\sigma\right] \delta u_\sigma \, dV - \int_S \left[\frac{\partial \psi}{\partial \eta_{\sigma\rho}} \nu_\rho - \frac{1}{3}\frac{\partial \psi}{\partial \eta_{\rho\rho}} \nu_\sigma\right] \delta u_\sigma \, dS = 0,$$

where S denotes the boundary of the body and ν is the outward unit normal to S. Actually, the surface integral can be restricted to the part Σ_1 of the boundary, since the quantities $\delta u_\sigma = 0$ over Σ_2 by hypothesis. In view of the arbitrary character of the virtual displacement δu, the above equation is now seen to decompose into the two following differential conditions, namely,

$$\left(\frac{\partial \psi}{\partial \eta_{\sigma\rho}}\right)_\rho - \frac{1}{3}\left(\frac{\partial \psi}{\partial \eta_{\rho\rho}}\right)_\sigma = 0, \qquad \text{over } V, \tag{5}$$

$$\frac{\partial \psi}{\partial \eta_{\sigma\rho}} \nu_\rho - \frac{1}{3}\frac{\partial \psi}{\partial \eta_{\rho\rho}} \nu_\sigma = 0, \qquad \text{over } \Sigma_1. \tag{6}$$

3. *Equations of Flow.*—It is generally considered admissible to replace the dynamical equations of motion by the equilibrium conditions $\sigma_{\alpha\beta,\beta} = 0$ in treating the plasticity problem. Using these conditions and differentiating the last set of relations (1) we find $s_{\alpha\beta,\beta} = p_{,\alpha}$ where we have put $p = -\sigma_{ii}/3$ for brevity. One also usually assumes that the plastic material is incompressible, which is expressed by the relation $u_{\alpha,\alpha} = 0$. Combining these conditions and making use of the assumption of section 1 that the tensor s is an invariant of the η tensor, we can write

$$\frac{\partial s_{\alpha\beta}}{\partial \eta_{\sigma\rho}} \eta_{\sigma\rho,\beta} = p_{,\alpha}; \qquad u_{\alpha,\alpha} = 0. \tag{7}$$

These are four relations for the determination of the four quantities u_α and p. To take account of compressibility, one sometimes replaces the last equation in (7) by the corresponding equation derived from ordinary elasticity theory. This leads to the equations

$$\frac{\partial s_{\alpha\beta}}{\partial \eta_{\sigma\rho}} \eta_{\sigma\rho,\beta} = p_{,\alpha}; \qquad p = -\frac{1}{3}\frac{Eu_{\alpha,\alpha}}{1-2\nu}, \tag{8}$$

where E is Young's modulus and ν is Poisson's ratio for the material. Again, we have four unknown quantities u_α and p and four equations (8) for the determination of these quantities. When the quantities u_α and p are found from either (7) or (8), we can determine the $e_{\alpha\beta}$ and the $\eta_{\alpha\beta}$ from the first two sets of relations (1). Also, the components $s_{\alpha\beta}$ can be determined from the relations $s = s(\eta)$, and, finally, the stress components $\sigma_{\alpha\beta}$ can be obtained from the last set of equations (1).

4. *Prevention of Overdetermination.*—When we add (7) or (8) to the conditions (5) and (6), we shall in general have an *overdetermined system.*[3] *To prevent this situation we impose conditions which are necessary and sufficient for the nonexistence of such overdetermination.* Since equations (7) or (8) are already sufficient for the determination of the quantities u_α and p, under the assumption that the functions $s_{\alpha\beta}(\eta)$ are known, this means that (5) and (6) must add no new conditions to the conditions given by (7) or (8). In particular, (5) and (6) can put no additional *algebraic conditions* on the quantities $\eta_{\alpha\beta,\gamma}$ over those imposed by (7) or (8). We shall now determine the precise consequences of this requirement.

The quantities $\eta_{\alpha\beta,\gamma}$ satisfy the relations $\eta_{\alpha\alpha,\rho} = 0$ from the definition of η as a deviation tensor, and it is obvious that there will be, in general, no additional algebraic conditions on the components of this tensor. Now (7) or (8) can put no additional algebraic conditions on the quantities $\eta_{\alpha\beta,\gamma}$, since the first set of equations in (7) or (8) can always be satisfied by a proper choice of the $p_{,\alpha}$ for any selection of the $\eta_{\alpha\beta,\gamma}$ and the remaining equation in (7) or (8) does not involve these quantities. Hence it follows from the above discussion that equations (5) must be satisfied whenever $\eta_{\alpha\alpha,\beta} = 0$. But we can write

$$\left(\frac{\partial \psi}{\partial \eta_{\sigma\rho}}\right)_p = \frac{\partial^2 \psi}{\partial \eta_{\sigma\rho}\partial \eta_{\alpha\beta}} \eta_{\alpha\beta,\rho},$$

$$\left(\frac{\partial \psi}{\partial \eta_{\rho\rho}}\right)_\sigma = \frac{\partial^2 \psi}{\partial \eta_{\rho\rho}\partial \eta_{\alpha\beta}} \delta_{\sigma\rho}\, \eta_{\alpha\beta,\rho}.$$

Hence we have

$$\left[\frac{\partial^2\psi}{\partial\eta_{\sigma\rho}\partial\eta_{\alpha\beta}} - \frac{1}{3}\frac{\partial^2\psi}{\partial\eta_{\gamma\gamma}\partial\eta_{\alpha\beta}}\delta_{\sigma\rho}\right]\eta_{\alpha\beta,\,\rho} = 0,$$

whenever $\eta_{\alpha\alpha,\,\rho} = 0$. From symmetry considerations we are led to express this condition as follows:

$$\left[\frac{\partial^2\psi}{\partial\eta_{\sigma\rho}\partial\eta_{\alpha\beta}} - \frac{1}{3}\frac{\partial^2\psi}{\partial\eta_{\gamma\gamma}\partial\eta_{\alpha\beta}}\delta_{\sigma\rho} - \frac{1}{3}\frac{\partial^2\psi}{\partial\eta_{\gamma\gamma}\partial\eta_{\sigma\rho}}\delta_{\alpha\beta}\right]\eta_{\alpha\beta,\,\rho} = 0,$$

whenever $\eta_{\alpha\alpha,\,\rho} = 0$.

For brevity let us denote the above symmetrical bracketed expression by $< \sigma\rho\alpha\beta >$, so that our condition becomes

$$< \sigma\rho\alpha\beta > \eta_{\alpha\beta,\,\rho} = 0, \qquad \text{whenever } \eta_{\alpha\alpha,\,\rho} = 0.$$

But we can also write

$$[< \sigma\rho\alpha\beta > - A_{\sigma\rho}\delta_{\alpha\beta}]\eta_{\alpha\beta,\,\rho} = 0, \qquad \text{whenever } \eta_{\alpha\alpha,\,\rho} = 0, \tag{9}$$

with arbitrary quantities $A_{\sigma\rho}$. Now three of these quantities η, e.g., the three components $\eta_{11,1}$, $\eta_{11,2}$, and $\eta_{11,3}$, can be selected as dependent, in that they can be considered to be determined by the relations $\eta_{\alpha\alpha,\,\rho} = 0$, after all other components $\eta_{\alpha\beta,\,\gamma}$ have been selected arbitrarily. Choose the A's in the above relations (9) by the condition $A_{\sigma\rho} = < \sigma\rho\,11>$. Then the coefficients of all dependent quantities $\eta_{\alpha\beta,\,\rho}$ vanish in the relations (9). The coefficients of the remaining quantities $\eta_{\alpha\beta,\,\rho}$ must therefore likewise vanish, owing to the arbitrary character of these quantities. Hence we have

$$< \sigma\rho\alpha\beta > = A_{\sigma\rho}\delta_{\alpha\beta}. \tag{10}$$

But from symmetry $< \sigma\rho\alpha\beta > = <\alpha\beta\sigma\rho>$. Hence, from (10), it follows that

$$A_{\sigma\rho}\delta_{\alpha\beta} = A_{\alpha\beta}\delta_{\sigma\rho}.$$

Putting $\alpha = \beta$ in these relations and summing on the repeated index, we now obtain

$$A_{\sigma\rho} = \frac{1}{3}A_{ii}\delta_{\sigma\rho} = B\delta_{\sigma\rho}.$$

Hence (10) becomes

$$< \sigma\rho\alpha\beta > = B\delta_{\sigma\rho}\delta_{\alpha\beta}. \tag{11}$$

In the treatment of relations (11), we have found it expedient to avail ourselves of the simplification afforded by the use of *canonical co-ordinates*. In the following section we develop the necessary formulas of differentiation relative to these co-ordinates.

5. *Canonical Co-ordinates.*—Consider the set of three equations

$$(\eta_{\alpha\beta} - \tau_i\delta_{\alpha\beta})h_i^{\ \beta} = 0, \qquad i \text{ not summed.} \tag{12}$$

The quantities τ_i in these equations are the principal values of the tensor η and are given as solutions of the equation formed by equating to zero the determinant of

the coefficients of the $h_i{}^\beta$. Since this determinant is of the third order, there will be three quantities τ_i, which in general will have distinct values; these values can be shown to be real. *We assume this general case in the following discussion.* Then the three vectors h_i which are determined as solutions of (12) will be unique to within algebraic sign and can be shown to be mutually orthogonal, i.e.,

$$h_i{}^\alpha h_i{}^\beta = \delta^{\alpha\beta}, \qquad \text{or} \qquad h_i{}^\alpha h_j{}^\alpha = \delta_{ij}. \tag{13}$$

We may suppose the designation of the vectors h_i to be such that the vector triad h_1, h_2, h_3 has the same orientation as the co-ordinate axes x^1, x^2, x^3. The rectangular co-ordinate system y having its origin at any point P and such that its y^1, y^2, y^3 axes coincide with the directions h_1, h_2, h_3 at P will be related to the underlying x system by a *proper* orthogonal transformation. Such a co-ordinate system y will be called *canonical.* An obvious property of the canonical system is that relative to this system and at its origin we have $\eta_{\alpha\beta} = \tau_\alpha \delta_{\alpha\beta}$. In fact, this relation follows immediately by considering equation (12) relative to canonical co-ordinates, since $h_i{}^\beta = \delta_i{}^\beta$ at the origin of the canonical system.

Multiplying (12) by $h_j{}^\alpha$ and summing on the repeated index α, we see that these equations can be written in the form

$$\tau_i \delta_{ij} = \eta_{\alpha\beta} h_i{}^\alpha h_j{}^\beta. \tag{14}$$

We now differentiate relations (13) and (14) partially with respect to the $\eta_{\sigma\rho}$. In carrying out this differentiation, it is advantageous from the formal standpoint to suppose that the quantities $\eta_{\sigma\rho}$ are completely independent and to take account of such properties as the symmetry of the components $\eta_{\sigma\rho}$ and the fact that the actual tensor η has a vanishing trace after all differentiations have been completed. Hence, from the second set of relations (13), we obtain

$$\frac{\partial h_i{}^\alpha}{\partial \eta_{\sigma\rho}} h_j{}^\alpha + h_i{}^\alpha \frac{\partial h_j{}^\alpha}{\partial \eta_{\sigma\rho}} = 0. \tag{15}$$

Similarly, by differentiation of (14), we are led to the relations

$$\begin{aligned} \frac{\partial \tau_i}{\partial \eta_{\sigma\rho}} \delta_{ij} &= \delta_{\alpha\sigma}\delta_{\beta\rho} h_i{}^\alpha h_j{}^\beta + \eta_{\alpha\beta} \frac{\partial h_i{}^\alpha}{\partial \eta_{\sigma\rho}} h_j{}^\beta + h_i{}^\alpha \frac{\partial h_j{}^\beta}{\partial \eta_{\sigma\rho}} \\ &= h_i{}^\sigma h_j{}^\rho + \tau_j h_j{}^\alpha \frac{\partial h_i{}^\alpha}{\partial \eta_{\sigma\rho}} + \tau_i h_i{}^\beta \frac{\partial h_j{}^\beta}{\partial \eta_{\sigma\rho}}, \end{aligned}$$

when use is made of (12). But on account of (15) these latter relations become

$$\frac{\partial \tau_i}{\partial \eta_{\sigma\rho}} \delta_{ij} = h_i{}^\sigma h_j{}^\rho + (\tau_i - \tau_j) h_i{}^\alpha \frac{\partial h_j{}^\alpha}{\partial \eta_{\sigma\rho}}.$$

Hence, according as $i = j$ or $i \neq j$, we now have

$$\frac{\partial \tau_i}{\partial \eta_{\sigma\rho}} = h_i{}^\sigma h_i{}^\rho, \tag{16}$$

$$(\tau_i - \tau_j) h_i{}^\alpha \frac{\partial h_j{}^\alpha}{\partial \eta_{\sigma\rho}} + h_i{}^\sigma h_j{}^\rho = 0, \qquad i \neq j. \tag{17}$$

The above relations (15), (16), and (17) are valid in arbitrary rectangular co-ordinate systems. Transforming to canonical co-ordinates and evaluating at the origin of the canonical system, we obtain

$$\frac{\partial h_j{}^i}{\partial \eta_{\sigma\rho}} + \frac{\partial h_i{}^j}{\partial \eta_{\sigma\rho}} = 0, \tag{18}$$

$$\frac{\partial \tau_i}{\partial \eta_{\sigma\rho}} = \delta_i{}^\sigma \delta_i{}^\rho, \tag{19}$$

$$\frac{\partial h_j{}^i}{\partial \eta_{\sigma\rho}} = \frac{-\delta_i{}^\sigma \delta_j{}^\rho}{\tau_i - \tau_j}, \qquad i \neq j. \tag{20}$$

Equations (18) have the simple interpretation that the derivatives of the $h_j{}^i$ are skew-symmetric in the indices i and j. Hence, in particular, if these indices have equal values, the derivative must vanish.

6. *Condition on ψ.*—The scalar invariant $\psi(\eta)$ of the η tensor can be regarded as a function $\psi(\tau)$ of the three principal values τ_1, τ_2, τ_3 of this tensor. This is seen immediately by transforming the invariant $\psi(\eta)$ to canonical co-ordinates and then using the fact that $\eta_{\alpha\beta} = \tau_\alpha \delta_{\alpha\beta}$ at the origin of the canonical system. Hence, by differentiation, we have

$$\frac{\partial \psi}{\partial \eta_{\sigma\rho}} = \frac{\partial \psi}{\partial \tau_i}\frac{\partial \tau_i}{\partial \eta_{\sigma\rho}} = \frac{\partial \psi}{\partial \tau_i} h_i{}^\sigma h_i{}^\rho, \tag{21}$$

when use is made of (16). There is a summation on the index i, of course, in these relations. Again differentiating (21), we find that

$$\frac{\partial^2 \psi}{\partial \eta_{\sigma\rho}\partial \eta_{\alpha\beta}} = \frac{\partial^2 \psi}{\partial \tau_i \partial \tau_j} h_i{}^\sigma h_i{}^\rho\, h_j{}^\alpha\, h_j{}^\beta + \frac{\partial \psi}{\partial \tau_i}\frac{\partial h_i{}^\sigma}{\partial \eta_{\alpha\beta}} h_i{}^\rho + \frac{\partial \psi}{\partial \tau_i} h_i{}^\sigma \frac{\partial h_i{}^\rho}{\partial \eta_{\alpha\beta}}.$$

At the origin of canonical co-ordinates these equations become

$$\frac{\partial^2 \psi}{\partial \eta_{\sigma\rho}\partial \eta_{\alpha\beta}} = \frac{\partial^2 \psi}{\partial \tau_\alpha \partial \tau_\sigma}\delta_{\alpha\beta}\delta_{\sigma\rho} + \left(\frac{\partial \psi}{\partial \tau_\rho} - \frac{\partial \psi}{\partial \tau_\sigma}\right)\frac{\partial h_\rho{}^\sigma}{\partial \eta_{\alpha\beta}},$$

as we see immediately when we make the substitution $h_i{}^\alpha = \delta_i{}^\alpha$ and use relations (18). Hence, in particular, we can write

$$\frac{\partial^2 \psi}{\partial \eta_{\sigma\rho}\partial \eta_{\alpha\beta}} = \left(\frac{\partial \psi}{\partial \tau_\rho} - \frac{\partial \psi}{\partial \tau_\sigma}\right)\frac{\partial h_\rho{}^\sigma}{\partial \eta_{\alpha\beta}}, \qquad \alpha \neq \beta; \qquad \sigma \neq \rho, \tag{22}$$

at the origin of the canonical system.

Let us now consider the invariant relations (11) in canonical co-ordinates, and let us then evaluate these relations at the origin of this system. Choosing $\sigma \neq \rho$ and $\alpha \neq \beta$, we then have

$$< \sigma\rho\alpha\beta > = \frac{\partial^2 \psi}{\partial \eta_{\sigma\rho}\partial \eta_{\alpha\beta}} = 0, \qquad \alpha \neq \beta; \qquad \sigma \neq \rho. \tag{23}$$

Comparison of (22) and (23) now yields

$$\left(\frac{\partial\psi}{\partial\tau_\rho} - \frac{\partial\psi}{\partial\tau_\sigma}\right)\frac{\partial h_\rho{}^\sigma}{\partial\eta_{\alpha\beta}} = 0, \qquad \alpha \neq \beta; \qquad \sigma \neq \rho,$$

or

$$\left(\frac{\partial\psi}{\partial\tau_\rho} - \frac{\partial\psi}{\partial\tau_\sigma}\right)\frac{\delta_\sigma{}^\alpha\delta_\rho{}^\beta}{\tau_\sigma - \tau_\rho} = 0, \qquad \alpha \neq \beta; \qquad \sigma \neq \rho,$$

on account of (20). In these relations there is, of course, no summation on repeated indices, and $\tau_\sigma \neq \tau_\rho$, since $\sigma \neq \rho$ and the τ's are distinct by hypothesis. Hence, taking $\alpha = \sigma$ and $\beta = \rho$ without summation, it follows that $\partial\psi/\partial\tau_\alpha = \partial\psi/\partial\tau_\beta$ for $\alpha \neq \beta$. In other words, we must have

$$\frac{\partial\psi}{\partial\tau_1} = \frac{\partial\psi}{\partial\tau_2} = \frac{\partial\psi}{\partial\tau_3}. \tag{24}$$

While these relations have been established at the origin of canonical co-ordinates by the above process, it is clear that they actually hold without regard to the co-ordinate system, on account of their invariant character.

From (21) and (24), we now have

$$\frac{\partial\psi}{\partial\eta_{\sigma\rho}} = \frac{\partial\psi}{\partial\tau_1}\sum_i h_i{}^\sigma h_i{}^\rho = \frac{\partial\psi}{\partial\tau_1}\delta_{\sigma\rho},$$

or

$$\frac{\partial\psi}{\partial\eta_{\alpha\alpha}} = \frac{\partial\psi}{\partial\tau_1}, \qquad \alpha \text{ not summed}, \tag{25}$$

$$\frac{\partial\psi}{\partial\eta_{\alpha\beta}} = 0, \qquad \alpha \neq \beta. \tag{26}$$

It is immediately seen from (25) and (26) that the invariant ψ is a function of the trace $\Sigma\eta_{\alpha\alpha}$ of the η tensor. A simple formal demonstration of this fact can be given as follows: Observing that the function $\psi(\eta)$ does not depend on the components $\eta_{\alpha\beta}$ for $\alpha \neq \beta$ on account of (26), let us make the substitution

$$\eta_{11} = \bar{\eta}_{11}; \qquad \eta_{22} = \bar{\eta}_{22}; \qquad \eta_{33} = \bar{\eta}_{33} - \bar{\eta}_{11} - \bar{\eta}_{22}. \tag{27}$$

Suppose that $\psi(\eta)$ becomes $\bar{\psi}(\bar{\eta})$ as the result of this substitution. Then, taking account of (25), we have

$$\frac{\partial\bar{\psi}}{\partial\bar{\eta}_{11}} = \frac{\partial\psi}{\partial\eta_{11}} - \frac{\partial\psi}{\partial\eta_{33}} = 0; \qquad \frac{\partial\bar{\psi}}{\partial\bar{\eta}_{22}} = \frac{\partial\psi}{\partial\eta_{22}} - \frac{\partial\psi}{\partial\eta_{33}} = 0.$$

Hence $\bar{\psi}$ depends only on the single quantity $\bar{\eta}_{33}$, which, from (27), is seen to be the trace of the η tensor.

Denoting by θ the trace of the tensor η, the above result can be expressed by saying that ψ is a function of θ alone. Hence

$$\frac{\partial\psi}{\partial\eta_{\sigma\rho}} = \psi'(\theta)\,\delta_{\sigma\rho}; \qquad \frac{\partial\psi}{\partial\eta_{\rho\rho}} = 3\,\psi'(\theta),$$

in which the prime denotes differentiation with respect to θ. Also,

$$\left(\frac{\partial\psi}{\partial\eta_{\sigma\rho}}\right)_{\rho} = \psi''(\theta)\theta_{,\sigma}; \qquad \left(\frac{\partial\psi}{\partial\eta_{\rho\rho}}\right)_{\sigma} = 3\psi''(\theta)\theta_{,\sigma}.$$

Hence conditions (5) are satisfied by the function $\psi = \psi(\theta)$. Turning now to the boundary relations (6), which were not used in the determination of the above form of the function ψ, we have

$$\frac{\partial\psi}{\partial\eta_{\sigma\rho}}\nu_{\rho} = \psi'(\theta)\nu_{\sigma}; \qquad \frac{\partial\psi}{\partial\eta_{\rho\rho}}\nu_{\sigma} = 3\psi'(\theta)\nu_{\sigma}.$$

It follows that the function $\psi = \psi(\theta)$ *automatically satisfies conditions (6).* In other words, $\psi(\theta)$ represents the most general form of the function ψ satisfying the variational condition (2) under the above hypotheses.

7. *The Yield Condition.*—If we now take cognizance of the fact that η is a deviation tensor, so that its trace $\theta \equiv 0$, it follows that $\psi = \text{const.}$ *In other words, the function* $\psi(s, \eta)$ *used originally in our variational problem must actually reduce to a constant as a function of the components* $\eta_{\alpha\beta}$. This result is of the nature of a *yield condition.* For illustration, suppose that $\psi = s_{\alpha\beta}s_{\alpha\beta}$ and that the relation between the tensors s and η is given by $s_{\alpha\beta} = \phi\eta_{\alpha\beta}$, where the proportionality factor ϕ is here to be thought of as a scalar invariant of the tensor η. Then, from the above result, we have the von Mises or quadratic yield condition $s_{\alpha\beta}s_{\alpha\beta} = 2k^2$, where k is a constant depending on the material properties of the medium. Hence $\phi^2\eta_{\alpha\beta}\eta_{\alpha\beta} = 2k^2$, or $\phi = \sqrt{2}\,k/\sqrt{\eta_{ij}\eta_{ij}}$, and hence

$$s_{\alpha\beta} = \frac{\sqrt{2}k\eta_{\alpha\beta}}{\sqrt{\eta_{ij}\eta_{ij}}}. \tag{28}$$

This relation between the tensor s and η is precisely that used by von Mises in his theory of plasticity.

8. *General Relation between the s and* η *Tensors.*—It has been shown by Erickson and Rivlin[4] that if the tensor s in an invariant of the tensor η, we must have

$$s_{\alpha\beta} = F\delta_{\alpha\beta} + J\eta_{\alpha\beta} + K\eta_{\alpha\sigma}\eta_{\sigma\beta}$$

for the three-dimensional case under consideration, where *the quantities* F, J, *and* K *are scalar invariants of the* η *tensor with respect to orthogonal transformations.* Imposing the condition that the trace of the tensors s and η must vanish, it follows that $3F = -K\eta_{ij}\eta_{ij}$. Using this relation to eliminate the scalar F from the above equations, we thus obtain

$$s_{\alpha\beta} = J\eta_{\alpha\beta} + K\eta_{\alpha\sigma}\eta_{\alpha\beta} - {}^1/_3K(\eta_{ij}\eta_{ij})\delta_{\alpha\beta}. \tag{29}$$

Selection of the function $\psi(s, \eta)$ will now lead to a relation between the scalars J and K. Thus, taking $\psi = s_{\alpha\beta}s_{\alpha\beta}$, we have

$$\eta_{ij}\eta_{ij}J^2 + 2JK\,\eta_{\alpha\sigma}\eta_{\sigma\beta}\eta_{\alpha\beta} + K^2[\eta_{\sigma\alpha}\eta_{\sigma\beta}\eta_{\alpha\tau}\eta_{\tau\beta} - {}^1/_3(\eta_{ij}\eta_{ij})^2] = \text{const.}$$

Having recourse to canonical co-ordinates, we see that we can now write

$$\tau_{(1)} = \eta_{ii} = \Sigma\tau_i; \qquad \tau_{(2)} = \eta_{ij}\eta_{ij} = \Sigma\tau_i\tau_i,$$
$$\tau_{(3)} = \eta_{\alpha\sigma}\eta_{\sigma\beta}\eta_{\alpha\beta} = \Sigma\tau_i\tau_i\tau_i,$$
$$\tau_{(4)} = \eta_{\alpha\sigma}\eta_{\sigma\beta}\eta_{\alpha\tau}\eta_{\tau\beta} = \Sigma\tau_i\tau_i\tau_i\tau_i.$$

These relations define the scalar invariants $\tau_{(1)}$, $\tau_{(2)}$, $\tau_{(3)}$, and $\tau_{(4)}$ in terms of the components $\eta_{\alpha\beta}$ and also give their values in terms of the principal values τ_i of the η tensor. Now we can easily establish the following general identity:

$$\tau_{(4)} = \tau_{(2)}^2 + 4\tau_1\tau_2\tau_3\tau_{(1)} - \frac{1}{2}[\tau_{(2)} - \tau_{(1)}^2]^2.$$

But $\tau_{(1)} = 0$ from the special character of η as a deviation tensor in our problem, and hence the above identity reduces to $\tau_{(4)} = \tau_{(2)}^2/2$. Using this result and putting $L = K/J$, the above relation between J and K can now be written as

$$\left[\tau_{(2)} + 2\tau_{(3)}L + \frac{1}{6}\tau_{(2)}^2L^2\right]J^2 = 2k^2,$$

where k is a material constant. We thus have

$$J = \frac{\sqrt{2}k}{\sqrt{\tau_{(2)} + 2\tau_{(3)}L + \frac{1}{6}\tau_{(2)}^2L^2}}, \qquad K = \frac{\sqrt{2}kL}{\sqrt{\tau_{(2)} + 2\tau_{(3)}L + \frac{1}{6}\tau_{(2)}^2L^2}}.$$

Hence the stress-strain relation (29) involves, in addition to the usual constant k, an *invariant of the material* in the form of the scalar function L. For $L = 0$, we have $K = 0$, and (29) reduces to the usual stress-strain relation (28).

[1] We use the summation convention in accordance with which an index which appears more than once in a term is to be summed over the values 1, 2, 3. Several exceptions to this rule occur, but these exceptions should be clear from the context. However, to prevent possible confusion in some of the formulas, or for additional emphasis, the summation sign, Σ, has been used in certain instances.

[2] Dr. H. Trent has suggested that the plastic region should be considered to be surrounded by an elastic region, since this would more closely conform with the actual situation which occurs when the plastic state is established as the result of applied stresses or displacements over the boundary of the body. We agree, of course, with this point of view. Actually, it might be advisable to consider also the possibility of a departure from the strict linearity of the relationships of the classical elasticity theory, since such linearity does not appear to persist up to the beginning of plastic flow. In neglecting these and other factors which might be mentioned, it is possible that we have been guilty of oversimplification of the problem. However, it is likely that the analytical procedures here employed can be incorporated in the treatment of such modifications of the point of view which we have adopted.

[3] The definition of the overdetermined system which is here adopted is the usual one, namely, that overdetermination occurs when the number of equations exceeds the number of dependent or unknown functions. It is evident that *in general* overdetermination will result in inconsistency in some sense or in incompatibility with the required boundary conditions of the problem. An example of such incompatibility with boundary conditions is found when one attempts to solve the shock-wave problem which arises when an airfoil is placed in a uniform supersonic stream and it is assumed that the flow is everywhere irrotational or isentropic.

[4] R. S. Rivlin and J. L. Ericksen, "Stress-Deformation Relations for Isotropic Materials." This paper will probably appear in *J. Rat. Mech. Analysis*.

NOTE ON THE DISTRIBUTION OF ZEROS OF EXTREMAL POLYNOMIALS

By J. L. Walsh and J. P. Evans

DEPARTMENTS OF MATHEMATICS, HARVARD UNIVERSITY AND WELLESLEY COLLEGE

Communicated March 12, 1954

If E is a closed bounded point set of the z-plane, the relation

$$\lim_{n \to \infty} [\max |p_n(z)|, z \text{ on } E]^{1/n} = \tau(E), \tag{1}$$

where $\tau(E)$ is the transfinite diameter of E, is satisfied[1, 2, 3] by numerous sequences of polynomials $p_n(z) \equiv z^n + a_1^{(n)} z + \ldots + a_n^{(n)}$ of least norm, whether norm is measured by the Tchebycheff method or by the line or surface integral of $|p_n(z)|^p$, $p > 0$, with weight function, or by numerous other methods with only mildly restrictive properties. If the zeros of the $p_n(z)$ lie on E or on some closed bounded set E' (containing E) whose complement K' is connected, equation (1) is known[3] to imply

$$\lim_{n \to \infty} |p_n(z)|^{1/n} = |\phi(z)| \tag{2}$$

for z in K', uniformly on any closed bounded subset of K', where $w = \phi(z)$ maps conformally but not necessarily one to one the complement K (assumed to be connected) of E onto the region $|w| > \tau(E)$ with $\phi(\infty) = \infty$. In other words, if $G(z)$ is Green's function in the extended sense for K with pole at infinity, and if the conjugate of $G(z)$ in K is $H(z)$, then $\phi(z) \equiv \exp [G(z) + iH(z) + \log \tau(E)]$. We note that $\lim_{|z| \to \infty} [G(z) - \log |z|] = -\log \tau(E)$. The term *extended* is used here (and similarly below) to indicate that $G(z)$ is defined as the limit of a sequence of Green's functions in the classical sense for a monotonic increasing sequence of regions in K which exhaust K, each region of the sequence bounded by a finite number of Jordan curves; we do not assume that $G(z)$ approaches zero as z approaches an arbitrary boundary point of K.

The question naturally arises as to a substitute relation for equation (2) in regions of the plane, in $K - K'$ or elsewhere, containing limit points of zeros of the $p_n(z)$, for equation (2) obviously cannot hold uniformly in such regions. As a response to this question, we consider *exact harmonic majorants* as a consequence of equation (1), and we shall obtain both an analogue of equation (2) and some results on the asymptotic distribution of the zeros of the $p_n(z)$. We prove the following theorem:

Theorem. *Let E be a closed bounded point set of the z-plane with positive transfinite diameter whose complement K is connected. Let $p_n(z) \equiv z^n + a_1^{(n)} z^{n-1} + \ldots + a_n^{(n)}$ be a sequence of polynomials satisfying equation (1). Then, throughout K with the point at infinity deleted, the function $G(z) + \log \tau(E) \equiv \log |\phi(z)|$ is an exact harmonic majorant for the sequence $[p_n(z)]^{1/n}$ and for every subsequence.*

Any closed bounded subregion D interior to K contains a number, d_n, of zeros of $p_n(z)$ satisfying

$$\lim_{n \to \infty} \frac{d_n}{n} = 0; \tag{3}$$

this same relation holds in any limited region D bounded by an analytic Jordan curve C in K containing in its interior only a subset of E on which G(z) has a removable singularity.

If C is any analytic Jordan curve in K containing the number c_n of zeros of $p_n(z)$ in its interior, we have

$$\lim_{n \to \infty} \frac{c_n}{n} = \frac{1}{2\pi} \int_C \frac{\partial G}{\partial \nu} ds = \omega(\infty, B, K), \tag{4}$$

where ν represents the exterior normal for C, and the integral is equal to the harmonic measure in the extended sense at infinity with respect to K of the part B of the boundary of K interior to C.

The polynomials $p_n(z)$ need not be defined for every n but merely for an infinite sequence of indices.

The statement that $\log |\phi(z)|$ is a *harmonic majorant* of the sequence $[p_n(z)]^{1/n}$ in $K - \infty$ is defined as requiring

$$\limsup_{n \to \infty} [\max |p_n(z)|^{1/n}, z \text{ on } Q] \leqq [\max |\phi(z)|, z \text{ on } Q] \tag{5}$$

for every bounded continuum Q (not a single point) in $K - \infty$. The concept of *exact harmonic majorant* requires

$$\limsup_{n \to \infty} [\max |p_n(z)|^{1/n}, z \text{ on } Q] = [\max |\phi(z)|, z \text{ on } Q] \tag{6}$$

for every such Q. If $\log |\phi(z)|$ is an exact harmonic majorant for $[p_n(z)]^{1/n}$ and for every subsequence, equation (6) becomes

$$\lim_{n \to \infty} [\max |p_n(z)|^{1/n}, z \text{ on } Q] = [\max |\phi(z)|, z \text{ on } Q],$$

for every such Q.

A special case of equation (1) is that of the Tchebycheff polynomials $t_n(z) \equiv z^n + t_1^{(n)} z^{n-1} + \ldots + t_n^{(n)}$ on E, namely, of least norm $[\max |t_n(z)|, z \text{ on } E]$; for these polynomials we have[2]

$$\lim_{n \to \infty} [\max |t_n(z)|, z \text{ on } E]^{1/n} = \tau(E), \tag{7}$$

and the zeros of $t_n(z)$ are known (Fejér) to lie in the convex hull $H(E)$ of E.

We denote generically by E_R $(R > 1)$ the boundary of the open subset K_R of K on which $G(z) > \log R$; thus E_R is closed and consists of the locus $G(z) = \log R$ in K, plus whatever points of E are limits of points z in K at which $G(z) > \log R$. If K is regular or if K is not regular and R is sufficiently large, E_R consists of a finite number of Jordan curves in K which are mutually exterior except perhaps for a finite number of points each belonging to several of the Jordan curves. We call K_R the *exterior* of E_R and call the complement, $\bar{E}_R$, of K_R the *closed interior* of E_R. We note that if $G(z) - \log R$ is Green's function in the classical sense for K_R, then the transfinite diameter of $\bar{E}_R$ is $\tau(\bar{E}_R) = R \cdot \tau(E)$.

These preliminaries complete, we now proceed with the proof of the theorem. By the generalized Bernstein lemma in extended form,[4] we have for every R (>1)

$$[\max |p_n(z)|, z \text{ on } \bar{E}_R] \leqq [\max |p_n(z)|, z \text{ on } E] \cdot R^n. \tag{8}$$

If Q is a bounded continuum in K and if R (>1) is the least number such that Q lies in $\bar{E}_R$, we have, by inequality (8),

$$[\max |p_n(z)|, z \text{ on } Q] \leqq [\max |p_n(z)|, z \text{ on } E] \cdot R^n,$$

which implies inequality (5) by virtue of our hypothesis (1) and the equation $\log |\phi(z)| = \log \tau(E) + \log R$ on E_R in K. Thus $\log |\phi(z)|$ is a harmonic majorant in the region $K - \infty$ for the sequence $[p_n(z)]^{1/n}$.

If $\log |\phi(z)|$ is not an exact harmonic majorant in $K - \infty$ for every subsequence of $[p_n(z)]^{1/n}$, we reach a contradiction. Some subsequence $[p_{n_k}(z)]^{1/n_k}$ satisfies the relation

$$\limsup_{n_k \to \infty} [\max |p_{n_k}(z)|^{1/n_k}, z \text{ on } Q] < [\max |\phi(z)|, z \text{ on } Q] \tag{9}$$

for some continuum Q (not a single point) in $K - \infty$ and hence[5] satisfies inequality (9) for every continuum Q in $K - \infty$. We choose Q as the locus E_R, where R (>1) is chosen so large that E_R is a single Jordan curve (hence a continuum) and so large that E lies interior to E_R; of course, Green's function with pole at infinity for the exterior of E_R in the classical sense is $G(z) - \log R$. Thus we have

$$\limsup_{n_k \to \infty} [\max |p_{n_k}(z)|^{1/n_k}, z \text{ on } E_R] < [\max |\phi(z)|, z \text{ on } E_R] = R \cdot \tau(E). \tag{10}$$

On the other hand, if $t_{n_k}(z)$ is the Tchebycheff polynomial of degree n_k *for the set* $\bar{E}_R$, we have, by the principle of maximum modulus and the definition of $t_{n_k}(z)$,

$$[\max |t_{n_k}(z)|, z \text{ on } \bar{E}_R] = [\max |t_{n_k}(z)|, z \text{ on } E_R] \leqq [\max |p_{n_k}(z)|, z \text{ on } E_R],$$

which, with equation (7) for the set $\bar{E}_R$ instead of E, using $\tau(\bar{E}_R) = R \cdot \tau(E)$, contradicts inequality (10). This contradiction completes the proof that $\log |\phi(z)|$ is an exact harmonic majorant in $K - \infty$ for $[p_n(z)]^{1/n}$ and for every subsequence.

Equation (3) follows for D in K at once[6] if D is a circle, and also for arbitrary D by covering D with a finite number of circles in K. We could alternately make use of the invariance under conformal mapping of the concept of least harmonic majorant, instead of covering an arbitrary D with circles.

Let C be an arbitrary analytic Jordan curve in K. There exists also an analytic Jordan curve Γ in K containing C in its interior but such that no point of E lies in the annular region between C and Γ. In the case that Γ is a circle, and hence by a conformal map whether Γ is a circle or not, the zeros of the $p_n(z)$ may be suppressed in a closed annular region, A, interior to Γ and in K, A containing C in its interior. We suppress the zeros of $p_n(z)$ by a method previously used[7] and replace $p_n(z)$ by $F_n(z)$; but we suppress only those zeros in A (not throughout the interior of C) and make essential use of equation (3), where D in equation (3) is chosen as A. To be explicit, $F_n(z)$ has precisely the same zeros interior to Γ as does $p_n(z)$, except that $F_n(z)$ is different from zero throughout A. The sequence $[F_n(z)]^{1/n}$, like $[p_n(z)]^{1/n}$, has $\log |\phi(z)|$ as an exact harmonic majorant in A and for every subsequence.

Whenever $u(z)$ is an exact harmonic majorant in a region B for a sequence $[\psi_n(z)]^{1/n}$ and for every subsequence, with $\psi_n(z) \neq 0$, it follows that, in any simply connected subregion B_1 of B, every subsequence of $[\psi_n(z)]^{1/n}$ admits a new subsequence which converges throughout B_1 to a function of modulus $e^{u(z)}$, uniformly on any closed subset of B_1.[8] Consequently, *we have* $\lim_{n \to \infty} |\psi_n(z)|^{1/n} = e^{u(z)}$ *uniformly on any*

closed subset of B. Since $\log|\varphi(z)|$ is an exact harmonic majorant in A for $[F_n(z)]^{1/n}$ and for every subsequence, with $F_n(z) \not\equiv 0$ in A, we have, throughout A,

$$\lim_{n\to\infty} |F_n(z)|^{1/n} = |\varphi(z)|, \tag{11}$$

uniformly on any closed subregion of A; the corresponding relation holds for the first partial derivatives of the logarithms of the two members of equation (11).

The number of zeros of $F_n(z)$ interior to C is given by

$$c_n - \delta_n = \frac{1}{2\pi}\int_C \frac{\partial \log|F_n(z)|}{\partial \nu}\, ds, \qquad \delta_n \geqq 0,$$

where $\delta_n/n \to 0$ as $n \to \infty$. Thus, from the relation of derivatives of logarithms corresponding to equation (11), we have

$$\lim_{n\to\infty}\frac{c_n}{n} = \lim_{n\to\infty}\frac{1}{2n\pi}\int_C \frac{\partial \log|F_n(z)|}{\partial \nu}\, ds = \frac{1}{2\pi}\int_C \frac{\partial G}{\partial \nu}\, ds. \tag{12}$$

The interpretation of this last member as harmonic measure in an extended sense, defined in terms of harmonic measure in the classical sense for a monotonic increasing sequence of subregions of K exhausting K, is to be published elsewhere.[3] Equation (12) includes both the case $\omega(\infty, B, K) > 0$ and the case $\omega(\infty, B, K) = 0$; the latter case occurs if C contains in its interior only a subset of E on which $G(z)$ has a removable singularity.

We mention the fact, which is new in the literature, that equation (2) holds in any subregion B' of K containing no zeros of the $p_n(z)$, uniformly on any closed subregion of B'.[8]

The theorem does not require that K be regular in the sense that the classical Green's function for K with pole at infinity exists. For instance, we may choose E as the set $|z| \leqq 1$ plus the point $z = 2$, $p_n(z) \equiv z^{n-1}(z-2)$, whence $[\max p_n(z), z \text{ on } E] = 3$, and equation (1) is satisfied.

A consequence of the theorem is that if C is any set consisting of a finite number of mutually exterior analytic Jordan curves in K, and if for regular K each point of E lies interior to some curve of C, and if for K not necessarily regular the totality of boundary points of K interior to the curves of C has harmonic measure (in the extended sence) unity at infinity, then the curves composing C contain in their interiors "almost all" the zeros of the polynomials $p_n(z)$; that is to say, the ratio to n of the number of zeros of $p_n(z)$ interior to these curves approaches unity as n becomes infinite.

The requirement in the theorem that $\tau(E)$ be positive cannot be omitted, even though E may contain an infinite number of points. As counterexample, we choose E as the set $z = 0, 1, 1/2, 1/4, \ldots, 1/2^n, \ldots$. For n odd ($n = 2m+1$) we choose

$$p_n(z) \equiv (z-2)^m (z-1)(z-1/2)\ldots(z-1/2^m),$$
$$[\max |p_n(z)|, z \text{ on } E] = |p_n(0)| = 1/2^{m(m-1)/2};$$

for n even ($n = 2m+2$) we choose

$$p_n(z) \equiv (z-2)^m (z-1)(z-1/2)\ldots(z-1/2^{m+1}),$$
$$[\max |p_n(z)|, z \text{ on } E] = |p_n(0)| = 1/2^{1+m(m+1)/2},$$

so, with $\tau(E) = 0$, equation (1) is satisfied. Yet the number of zeros of $p_n(z)$ in the point $z = 2$ not in E is m, with $m/n \to 1/2$, in contradiction to equation (3).

The classical polynomials of least norm on a set E have all their zeros on the convex hull $H(E)$ of E; nevertheless, our theorem makes no restriction on the boundedness of the zeros of $p_n(z)$. It may actually occur that the zeros of the totality of the $p_n(z)$ are everywhere dense in the plane. For instance, we choose E as $|z| \leqq 1$, $p_n(z) \equiv z^{n-1}(z - \alpha_n)$, where the α_n are the (countable) totality of points of form $(j + ik)/m$, with j, k, m, integers and the α_n so chosen that $|\alpha_n| + 1 \leqq n$; we have $[\max |p_n(z)|, z \text{ on } E] \leqq n$, so equation (1) is satisfied.

Let a point set E with rectifiable boundary C consist of a finite number of mutually disjoint Jordan arcs or closed Jordan regions, and let orthogonality be defined in terms of a line integral over C. Alternately, let E consist of a finite number of mutually exterior closed Jordan regions, and let orthogonality be defined in terms of a surface integral over E. Under very broad conditions[3] on the weight function, if $\{P_n(z)\}$ is the set of normal orthogonal polynomials of respective degrees n, the zeros of $P_n(z)$ lie in $H(E)$, and we have

$$\lim_{n \to \infty} |P_n(z)|^{1/n} = |\phi(z)|/\tau(E) \tag{13}$$

uniformly on any closed set in $K - H(E)$. If a function $f(z)$ is single-valued and analytic throughout the interior of E_ρ but not throughout the interior of any $E_{\rho'}$, $\rho' > \rho > 1$, the formal expansion

$$f(z) \equiv \sum_{n=0}^{\infty} a_n P_n(z) \tag{14}$$

converges[9] throughout the interior of E_ρ, uniformly on any closed set interior to E_ρ, but does not converge throughout any region exterior to E_ρ. Moreover, we have

$$\limsup_{n \to \infty} |a_n|^{1/n} = 1/\rho. \tag{15}$$

It is a consequence of our main theorem that *the series (14) cannot converge uniformly on any continuum (not a single point) exterior to* E_ρ. Indeed, the sequence $[P_n(z)]^{1/n}$ and every subsequence has $\log |\phi(z)| - \log \tau(E)$ as exact harmonic majorant in $K - \infty$; this follows from equation (6) for the orthogonal polynomials $p_n(z) \equiv z^n + \ldots$, where[3]

$$P_n(z) \equiv p_n(z)/[\text{norm of } p_n(z)], \quad \lim_{n \to \infty} [\text{norm of } p_n(z)]^{1/n} = \tau(E). \tag{16}$$

If (14) converges uniformly on a continuum Q (not a single point) exterior to E_ρ, we have, by equations (15) and (16),

$$\begin{aligned} 1 &\geqq \limsup_{n \to \infty} [\max |a_n P_n(z)|, \quad z \text{ on } Q]^{1/n} \\ &= \lim_{n \to \infty} [\max |p_n(z)|, \quad z \text{ on } Q]^{1/n}/\rho\tau(E), \end{aligned}$$

which contradicts equation (6), since the second member of equation (6) is greater than $\rho\tau(E)$.

The theory[10] applied in the present note has previously been used primarily in the study of sequences of functions approximating to a given analytic function, rather than in the study of the relations between various asymptotic properties of other sequences of functions. However, the method of proof of our main theorem is of wide applicability in the study of sets of functions other than polynomials, such as orthogonal functions defined by extremal properties in a region which is not necessarily simply connected. The present writers plan to return to this topic on another occasion.

The authors are informed that some results on the asymptotic distribution of zeros of Tchebycheff polynomials have recently been developed also by Dr. J. L. Ullman. The researches contained in the present note have been done independently.

[1] P. Davis and H. Pollak, "On an Equivalent Definition of the Transfinite Diameter," *Proc. Amer. Math. Soc.*, **3**, 152–155 (1952).

[2] M. Fekete, "Über die Verteilung der Wurzeln bei gewissen algebraischen Gleichungen mit ganzzahligen Koeffizienten," *Math. Zs.*, **17**, 228–249 (1923).

[3] M. Fekete and J. L. Walsh, "On the Asymptotic Behaviour of Polynomials with Extremal Properties, and of Their Zeros" (to appear).

[4] J. L. Walsh, *Interpolation and Approximation by Rational Functions in the Complex Domain* ("Colloquium Publications of the American Mathematical Society," Vol. **20** [1935]), pp. 77, 87.

[5] J. L. Walsh, "Overconvergence, Degree of Convergence, and Zeros of Sequences of Analytic Functions," *Duke Math. J.*, **13**, 195–234 (1946), Corollary to Theorem 1.

[6] *Ibid.*, Theorem 16.

[7] *Ibid.*, Theorem 15.

[8] *Ibid.*, p. 223, Remark 1.

[9] J. L. Walsh, *Interpolation and Approximation by Rational Functions in the Complex Domain*, secs. 6.6–6.7.

[10] Developed in J. L. Walsh, "Overconvergence, Degree of Convergence, and Zeros of Sequences of Analytic Functions."

THE EFFECTS OF CHORIOALLANTOIC TRANSPLANTS OF ADULT CHICKEN TISSUES ON HOMOLOGOUS TISSUES OF THE HOST CHICK EMBRYO

BY JAMES D. EBERT

BROOKHAVEN NATIONAL LABORATORY* AND INDIANA UNIVERSITY*

Communicated by B. H. Willier, March 15, 1954

Weiss has reported that the incorporation of organ fragments (e.g., liver or kidney) from 6-day-old chick embryos into the area vasculosa of 4-day-old chick embryo hosts results in a striking and relatively specific response by the homologous tissue of the host.[1] In the case of liver grafts, for example, after an initial decline, the weights of the hosts' livers steadily increase in comparison with those of controls, until, on the seventh postoperative day, livers of experimental hosts are over 40 per cent heavier than those of controls. In further tests of the organ-specific aspects of the effect, embryonic kidney brei was injected intravenously into younger embryos.[2] Mitotic counts after one or two days showed a 50 per cent increase above normal in the host kidneys, with high electivity of the effect. Weiss has sug-

gested that "either specifically shaped parts of liver cell proteins can act directly as nuclei for further synthesis when entering the appropriate environment, i.e., another liver cell, or they act merely as models for moulds which would turn out more of the original product." He has emphasized that, while the evidence at hand supports the general principle of organ-specific chemical growth regulation by homologous cell constituents, the mode of action is still a matter of conjecture; no selection has been possible between the "building-block" and "template" (catalytic) mechanisms. As a corollary of the theory of organ-specific growth regulation, it is argued that growth and differentiation, being somewhat antagonistic, may react in opposite directions. Hence, under conditions favoring growth, differentiation may be suppressed. Preliminary experiments have been cited in support of this view.[3, 4, 5]

The aim of the present investigation is to test critically the template hypothesis of growth by an immunochemical and biochemical analysis of the specificity, time of origin, and mechanism of action of the factors concerned in the enlargement of host chick embryonic organs following grafts of the homologous tissue from older donors. At the outset of the study the major experimental target was the chick spleen.

It has been established that transplants of adult chicken spleen made to the chorioallantoic membrane of 9-day-old host chick embryos elicit a relatively specific response on the part of the homologous tissue of the host embryo, viz., an increase in mass.[6, 7] The effect is class-specific and quantitatively tissue-specific.[8] Thus, although adult chicken spleen grafts result in a striking increase in weight of the host spleen, grafts of adult mouse spleen are ineffective. Moreover, of eleven adult chicken tissues tested in addition to spleen, only two, thymus and liver, affect the host spleen, and in each case the effect is quantitatively far less than that observed with splenic transplants. It has been shown, in addition, that the ability of chicken spleen grafts to effect host spleen enlargement is correlated directly with the age of the donor, the factor or factors concerned first appearing in the embryonic chick spleen at approximately the fourteenth day of development and increasing thereafter until the adult level is reached. The time of appearance of the specific factors concerned in host spleen enlargement is paralleled by the development of a group of three spleen-specific antigens, as revealed by immunochemical studies. These findings suggest a correlation between the progressive increase with age in the ability of a spleen graft to evoke enlargement of the host's spleen and the differentiation, during a similar period in ontogeny, of a specific group of spleen antigens.[8, 9]

The desirability of extending this line of research is apparent. Among the pertinent questions raised are the following: (1) Is the enlargement of the embryonic chick spleen following a homologous transplant due to an increase in protein, or is it the result of an accretion of smaller molecules? (2) If, as postulated earlier, the enlargement results from a transfer of protein or other tissue-specific molecules, what is the magnitude of the transfer? Is a small number of molecules involved (template or catalytic hypothesis), or is a larger population of molecules transferred? The following experiments were carried out in an attempt to answer these questions. A preliminary report[10] of the findings was made before the American Society of Zoölogists, December 28, 1953.

METHODS AND RESULTS

1. *Determination of Protein Content of Control and Stimulated Embryonic Spleens.*—Is the enlargement of the host embryo spleen following a graft of adult spleen the result of an increased protein content? The following analyses indicate that an affirmative answer to this question is justified. The approach has been to compare the effects on the total and nonprotein nitrogen content of host embryo spleens of the following operative procedures: (1) sham operations, in which the complete chorioallantoic transplantation procedure was carried out but only a drop of sterile Ringer-Locke solution was added; (2) chorioallantoic transplants of adult chicken spleen, inactivated by fixation with cold 95 per cent ethanol, followed by storage at −20° C.; (3) grafts of fresh adult chicken liver; (4) grafts of fresh adult chicken spleen. The chorioallantoic grafting technique employed has been described previously.[7, 8] Donor tissues were from adult fowls of the New Hampshire Red breed, of either sex. Host embryos were from eggs of the same breed of fowls, incubated for 9 days at 37°.5 C. Following the operations, eggs were returned to the incubator for 8 days (total host age at recovery, 17 days), after which the grafts were recovered and fixed in Allen's B-15 or Bouin's picroformol for subsequent histological study. The spleens of the host embryos were removed rapidly to ice-cold 0.15 *M* NaCl, phosphate-buffered at pH 7.0; each spleen was washed carefully in three changes of the saline solution and stored in 1 ml. of buffered saline at −20° C. A total of 308 operations was performed; 249 host spleens were recovered. Of this number, 110 individual spleens were weighed rapidly on a Roller-Smith balance and digested separately in the Kjeldahl reaction mixture recommended by Brüel, Holter, Linderstrøm-Lang, and Rozits.[11] The total nitrogen of each spleen was determined by the micro-Kjeldahl method described by Niederl and Niederl.[12] The results are summarized in Table 1.

TABLE 1

TOTAL NITROGEN OF CONTROL AND ENLARGED SPLEENS

No. Cases	Treatment	Mean Weight of Host Spleens (Mg.) ± Standard Error of Mean	Mean Total Nitrogen (μg/Mg Tissue [Wet]) ± Standard Error of Mean
27	Sham operations	11.2 ± 0.3	16.9 ± 1.0
27	Grafts of inactivated adult chicken spleen	9.7 ± 1.3	17.0 ± 0.8
25	Grafts of fresh adult chicken liver	21.0 ± 1.4	17.5 ± 0.5
31	Grafts of fresh adult chicken spleen	44.4 ± 2.5	16.4 ± 0 5

The remaining 139 spleens were analyzed for both total and nonprotein nitrogen. Spleens were not assayed individually but were pooled in groups of from four to six. Both wet and dry weights of each lot of splenic tissue were determined. Nonprotein nitrogen was determined as the Kjeldahl nitrogen present in an aliquot of a protein-free filtrate of tissue homogenate, the protein being removed by precipitation with trichloroacetic acid. The data are given in Table 2.

Since all the preceding determinations were made on spleens that had been stored at −20° C. for periods ranging from 1 to 17 weeks and thawed before analysis, thus being subject to possible leaching, it seemed advisable to perform a similar set of determinations on fresh material. For this purpose an experiment was set up in

which a total of 84 sham operations and chorioallantoic grafts was made. Eight days later, 73 host spleens were recovered rapidly, as previously described, except that digestions were started immediately. The data clearly substantiate the findings obtained with frozen tissues, namely, that the increase in weight of stimulated spleens is paralleled by an increased nitrogen content.

The mean weight of the host spleens stimulated by homologous grafts is approximately four times that of the controls. The total nitrogen content is similarly increased; on the basis of micrograms of nitrogen per milligram of tissue, there are no significant differences in the several groups. These findings clearly support the view that the increase in weight of host embryo spleens results from an increased protein content.

2. *Selective Incorporation of Radioactivity from Proteins of Labeled Chorioallantoic Grafts of Adult Chicken Tissues into Proteins of Homologous Tissues of Host Chick Embryos.*—Does enlargement of the host spleen result from a transfer of tissue-specific molecules from the transplant to the homologous host tissue? If so, what is the magnitude of the transfer? In attacking this problem, the basic approach involved the transplantation of adult chicken and mouse tissues, labeled with S^{35}, to

TABLE 2

TOTAL AND NONPROTEIN NITROGEN CONTENT OF CONTROL AND ENLARGED HOST SPLEENS

Series	Mean Total Wet Weight (Mg.)	Mean Total Dry Weight (Mg.)	Ratio: Wet Weight/ Dry Weight	Total N (μg/Mg Tissue [Dry])	Nonprotein N (μg/Mg Tissue [Dry])
Sham operated	63.6	11.0	5.8	105	2.9
Inactivated chicken spleen grafts	77.8	13.2	5.8	115	4.3
Fresh chicken liver grafts	102.2	17.3	5.9	106	2.9
Fresh chicken spleen grafts	262.4	43.6	6.0	108	3.2

the chorioallantoic membrane of host embryos. At the time of recovery, the grafts and several host tissues were analyzed to determine the extent of transfer and incorporation of radioactivity.

Adult New Hampshire Red fowls and adult white mice were injected intraperitoneally with S^{35}-labeled L-methionine (Abbott Laboratories). Animals were fasted from 1 to 4 hours prior to, and 12 hours following, the injection, at which time they were sacrificed by decapitation. The standard dosage of labeled methionine injected, 175 μc/kg, was determined in earlier studies, in which a total of 21 adult fowl were injected with a graded series of doses of labeled methionone (21.1–231.4 μc/kg). The objective was to achieve as high a level of activity as possible in the protein fraction of the tissues to be engrafted, without radiation damage. This is accomplished with approximately 175 μc/kg.[13] At sacrifice the following steps were carried out as rapidly as possible: (1) spleen, kidney, liver, gonad, and heart samples were fixed for histologic and radioautographic preparations; (2) spleen, kidney, and liver samples were weighed and stored in the Deepfreeze, pending digestion, for determination of specific activity; (3) the remainder of the spleen and kidney was weighed and used for transplants, which were started within 30 minutes after sacrifice of the donor.

All grafts were made to the chorioallantoic membranes of 9-day-old chick embryo hosts (New Hampshire Red breed). After transplantation, eggs were reincubated at approximately 37°5 C. for either 3 or 5 days, after which the transplants and the host spleen, liver, and kidney were recovered for analysis. One out of every five grafts, together with its homologous host organ, was fixed in Bouin's picroformol for histologic and radioautographic analysis. All other grafts and the host tissues indicated above were weighed rapidly on a Roller-Smith balance and plunged into ice-cold 0.15 *M* NaCl. Host spleens were pooled in sufficient numbers to weigh at least 50 mg. (wet weight). Host kidneys and livers were pooled in groups corresponding to each group of spleens, and equivalent aliquots were taken after homogenization. All tissues were homogenized and separated into trichloroacetic acid–insoluble and –soluble fractions. Each fraction was then digested to dryness in a mixture of fuming nitric and perchloric acids. The digestate was washed with 6 *N* HCl and again evaporated to dryness. Sulfate was precipitated from an aqueous solution of the digestate by addition of 4 per cent benzidine dihydrochloride in 0.4 *N* HCl and was deposited by filtration on a fritted glass disk. The specific

TABLE 3

INCORPORATION OF RADIOACTIVITY FROM PROTEINS OF LABELED SPLEEN GRAFTS

(Specific activity of host tissue protein/specific activity of graft tissue protein)

	EXPERIMENTAL SERIES			
	1	2	3	4
Host spleen	0.14	0.16	0.04	0.12
Host liver	.04	.07	.018	.05
Host kidney	0.04	0.06	0.016	0.04

TABLE 4

INCORPORATION OF RADIOACTIVITY FROM PROTEINS OF LABELED KIDNEY GRAFTS

(Specific activity of host tissue protein/specific activity of graft tissue protein)

	EXPERIMENTAL SERIES			
	1	2	3	4
Host spleen	0.038	0.06	0.025	0.05
Host liver	.038	.05	.031	.06
Host kidney	0.098	0.13	0.059	0.16

activity of the radiosulfur in benzidine sulfate precipitates of protein and nonprotein fractions of donor and host tissues was determined as counts per minute per micromole of sulfate. The extraction, digestion, and precipitation methods have been described by Forker *et al.*;[14] the filtration method for collection of precipitates has been described fully by Van Slyke *et al.*;[15] hence details of these procedures need not be given here. Radioactivity of the benzidine sulfate precipitates was counted in a methane flow proportional counter.[16] All data are corrected for background, decay, and self-absorption. Sulfate was determined by titration with NaOH, while hot, to the phenol red point.

The following results are based on analysis of donor and host tissues from a total of 902 transplants and injected eggs. The principal object of the study was the spleen. The choice of a second, or control, tissue for comparison lay between heart and kidney, both of which may be transplanted successfully. Adult kidney was selected because its specific activity was considerably higher than that of adult heart. The incorporation into host embryonic tissues of radioactivity from proteins of labeled spleen grafts and labeled kidney grafts, respectively, is shown in

Tables 3 and 4. The data are expressed as specific activity of host tissue protein/specific activity of graft protein; each value is the mean of two samples, determined in duplicate. The results are consistent. They clearly indicate a highly selective transfer of radioactivity from graft to homologous host tissue.

In addition, one series of experiments was performed comparing the incorporation of radioactivity from transplants of labeled chicken and mouse spleen. The donor chickens were injected with 175 μc of labeled methionine/kg of body weight; the donor mice received a somewhat higher dose, 225 μc/kg. The data are summarized in Table 5; each value is the mean of two samples, determined in duplicate. It will be noted that there is no indication of selective incorporation in the chick tissues of radioactivity from transplants of mouse spleen, a fact which corroborates earlier findings relative to the class-specificity of the effect.

TABLE 5

COMPARISON OF INCORPORATION OF RADIOACTIVITY FROM LABELED CHICKEN AND LABELED MOUSE SPLEEN GRAFTS

(Specific activity of host tissue protein/specific activity of graft tissue protein)

Graft	Host Spleen	Host Liver	Host Kidney
Labeled chicken spleen	0.033	0.012	0.013
Labeled mouse spleen	0.009	0.015	0.012

TABLE 6

INCORPORATION OF RADIOACTIVE METHIONINE VIA THE YOLK SAC

SPECIFIC ACTIVITIES OF HOST TISSUES (C.P.M./MICROMOLE OF SULFATE—CORRECTED)

Methionine Injected μc	Treatment before Injection	Spleen	Liver	Kidney
0.045	Sham operation	22.9	17.6	25.4
.045	Transplants nonradioactive spleen	26.6	22.6	24.6
.075	Sham operation	71.3	66.3	57.3
0.075	Transplants nonradioactive spleen	81.5	54.2	106

To summarize the findings thus far: (1) The results show clearly that, following transplantation of adult chicken spleen or kidney, the specific activity of the homologous host tissue is at least 2.4 times that of other embryonic tissues. The transfer is selective both from spleen to spleen and from kidney to kidney. (2) There is no evidence of selective incorporation of radioactivity following transplantation of labeled mouse spleen. (3) In addition, the magnitude of the transfer is greater than might be expected if only a template or catalytic mechanism were involved, for the data indicate that, as early as three days postoperatively, as much as 15 per cent of the host tissue protein may be derived, or contain specific molecular contributions, from the proteins of the graft.

To explore this possibility further, a second major group of experiments was performed. In these experiments, S^{35}-labeled methionine was injected into the yolk sac of 363 nine-day-old embryos immediately following either transplants of nonradioactive spleen or sham operations. Grafts and host tissues were recovered 3 days postoperatively. Although the host spleens were enlarged in the former group, and hence total activity was greater, the specific activities (activity per micromole of sulfate) in the tissues of the two groups did not differ significantly. The data are presented in Table 6. Each value is the mean of two experiments

performed in duplicate. Coupled with the previous findings, these results afford support to the view that the enlargement of embryonic host organs following homologous organ grafts from older donors is due at least in large part to a selective incorporation into the proteins of the host tissue of tissue-specific components from the transplant. Thus the question is raised as to the nature of the specific components transferred. Is it possible that whole cells of the graft are incorporated into the host tissues? Or does the transfer involve subcellular components, e.g., tissue-specific proteins or specific components of the protein molecule larger than amino acids? The data presented in the following section favor the latter view.

3. *Desoxyribose Nucleic Acid Content of Control and Stimulated Embryonic Spleens.*—Microscopic examination of sections of embryonic host kidney from the several control and experimental series offers no support for the first alternative, viz., that whole cells are transferred. The histology of the kidney is sufficiently straightforward so that a comparative study of control and experimental tissues offers little difficulty. Cell counts and measurements of cell size now in progress indicate that the major change in the embryonic host kidney 3 days after the transplantation of adult kidney is an increase in cell size. The mitotic count is increased only slightly. Such data, however, are not obtained readily for host embryo spleens; the spleen is not favorable material for reliable measurements of cell size. In order to obtain quantitative data concerning the possibility of cell transfer in the splenic series, therefore, the following approach has been adopted: It is assumed that if whole (nucleated) cells are transferred, then the desoxyribose nucleic acid (DNA) content of the host spleen should increase concomitantly with the increase in protein. If the transfer is only of protein or specific protein constituents, then the DNA content per milligram of spleen should be lower in the stimulated series than in the control series. It may be argued that this assumption is too great an oversimplification of a complex situation. Growth is, of course, a complex phenomenon which cannot be reduced to the measurement of a single parameter. Moreover, as Weiss has shown, there is an initial burst of mitotic activity in embryonic kidney following homologous grafts. If the splenic response to spleen grafts is comparable, then not only the possibility of transfer of whole cells but an increase in DNA due to stimulation of mitotic activity must be considered. If whole cells or nuclei are not transferred, but mitotic activity is increased, then it is possible that any difference may be obscured, or at least reduced. With these limitations in mind, the DNA content of spleens from the several control and experimental series has been determined. Host spleens are recovered 8 days postoperatively (host age, 17 days) following sham operations, control grafts, or spleen grafts; weighed rapidly; and plunged into ice-cold 0.04 *M* citric acid. Spleens are pooled within the various groups in sufficient numbers to make 100 mg. (wet weight) per determination. Spleens are homogenized, and the DNA extracted by the method described by Schneider.[17] DNA is determined by the diphenylamine reaction (as defined by Dische),[18] and expressed as micrograms DNA per milligram of tissue. Two hundred and forty-five fresh host spleens have been analyzed to date; the data are given in Table 7. These findings also are supported by analyses of 243 additional host spleens which had been stored at $-20°$ C. for varying periods prior to extraction. In both the fresh and the frozen group, it is clear that the DNA content of the enlarged spleens does not increase concomitantly with the increase in

protein, the difference between the DNA contents of control and enlarged spleens being highly significant. These findings are not compatible with the idea that significant numbers of whole cells are transferred from graft to host spleen;

TABLE 7

DNA CONTENT OF ENLARGED AND CONTROL HOST SPLEENS

Series	No. Cases	Mean Wet Weight of Host Spleens (Mg.) ± Standard Error of Mean	DNA of Host Spleens (μg/Mg) ± Standard Error of Mean
Sham operated	73	9.3 ± 0.3	6.7 ± 0.15
Adult chicken heart grafts	99	8.4 ± 0.3	6.5 ± .19
Adult chicken spleen grafts	73	37.1 ± 3.8	4.7 ± 0.21

together with the results discussed previously, they support the idea of a selective incorporation from grafts into homologous host tissues of tissue-specific proteins, or specific constituents of proteins larger than amino acids.

DISCUSSION

The problems of protein biosynthesis and embryogenesis are inextricably interwoven; yet it must be emphasized that the mechanism of protein biosynthesis is only one facet of the more complex question of the forces involved in cellular differentiation and growth. The term "differentiation" encompasses the totality of reactions of which a cell is capable at any given stage and includes the sum total of changes in cell characters from one stage to the next. Contrary to the implications of some definitions,[19] several lines of evidence may be interpreted as indicating that differentiation may occur through the loss or suppression of synthetic activity.[20]

As defined meaningfully by Weiss,[21] growth is the increase in that part of the molecular population of an organic system which is synthesized within that system. Growth is essentially synonymous with reproduction: "The fundamental feature is the capacity of a reproductive system to procreate more systems with similar reproductive faculty." This definition includes only the biosynthesis of proteins and other complex molecules of the self-reproducing type. Implicit in the definition is the fact that growth involves a net increase in protein; whether it results from an increased turnover of protein[22] or an inhibition of protein breakdown,[23] or both, is not clear.

The findings of the present study bear directly both on the specific problem of the mechanism of protein biosynthesis and on the more general question of the interrelationship of protein biosynthesis with differentiation and growth. Let us consider first the mechanism of protein biosynthesis: The processes which convert amino acids into proteins have been studied from several points of view, viz., intracellular site of synthesis, energy requirements, and rate of incorporation. The major experimental targets have been the initial combination of amino acids, the step termed "peptidization" by Lipmann,[24] and the role of possible "template" mechanisms in establishing protein specificity.[25] In addition, the possible role of precursor "type" proteins (urproteins or proteinogens) has been discussed.[26] Until recently, less attention has been paid to reactions of possible intermediates. Bor-

sook *et al.* have studied the composition of several large tissue peptides which may represent stages in protein biosynthesis.[27] Anfinsen and Steinberg have presented data on nonuniform labeling of different residues of the same amino acid in ovalbumin which favor the concept of peptide intermediates.[28] Moreover, from a tracer study of serum albumin synthesis, Peters has concluded that amino acids are in "intermediate compounds" for at least 20 minutes before the fully formed protein can be detected.[29] These intermediates do not resemble the protein in serological properties or electrophoretic or solubility behavior. Snellman and Danielsson have concluded from a study of biosynthesis of reserve globulins in pea seeds that the synthesis is a stepwise phenomenon, involving the formation of oligopeptide intermediates.[30] In addition to the several lines of evidence that favor a stepwise biosynthesis of protein involving peptide intermediates, one recent investigation has produced evidence which strongly suggests that embryonic cells may utilize a protein nutrient source without the release of free amino acids. Francis and Winnick have compared the mode of utilization by tissue cultures of cardiac muscle of free C^{14}-labeled amino acids with that of the same amino acids in protein-bound form.[31] From experiments in which relatively large concentrations of nonisotopic amino acids were added to the medium as metabolic traps, it was concluded that most of the C^{14} singly labeled proteins could be transferred from a nutrient embryo extract medium to heart muscle protein without the release of free amino acids. A similar conclusion was reached in experiments with doubly labeled proteins. Despite the convergence of opinion indicated in the several lines of investigation cited above, the facts available are not consistent enough to warrant any generalization. Muir, Neuberger, and Perrone, for instance, have presented data which indicate that there are no differences in specific activities of valine at terminal and nonterminal positions of hemoglobin synthesized by the rat supplied with the labeled amino acid.[32] Halvorson and Spiegelman have concluded that already formed proteins have no direct role in the formation of adaptive enzymes without passing through the free amino acid pool.[33]

The findings of the present investigation offer strong support for the conclusion of Francis and Winnick that embryonic cells can utilize tissue proteins without first breaking down the proteins into free amino acids, and, in fact, they favor the further conclusion that the incorporation is of a highly specific nature, e.g., that splenic proteins or specific constituents thereof are selectively incorporated by the developing spleen, a possibility which was not tested in the experiments by Francis and Winnick. It appears unlikely that the degree of electivity reported here could be accounted for on the basis of transfer of amino acids. The conclusion that the specific effect of grafts upon homologous host tissues is due to a selective incorporation from grafts into proteins of the host tissues of tissue-specific protein constituents larger than amino acids must be regarded as tentative, pending the outcome of further experiments. A consideration of the data, together with the findings of others, however, indicates this view to be acceptable as a working hypothesis, subject to further test.

It will be of interest to examine next the bearing of this working hypothesis on the concept of organ-specific growth control. The general principle of the control of of growth by organ-specific substances has been subjected to several experimental tests. The establishment of the class-specific and quantitative organ-specific as-

pects of the principle, and the conclusion that the increase in size of homologous tissues elicited by grafts results from a true increase in protein content, are based on a consideration of the data obtained in a sustained investigation involving, to date, analyses of graft and host tissues of 3,446 chorioallantoic transplants and variously injected and sham-operated embryos, employing the techniques of bio-, immuno-, and radiochemistry in conjunction with the more classical methods of experimental embryology. With respect to the mechanism of action of organ-specific growth-controlling substances, the working hypothesis suggested above permits us to focus on several clear experimental targets. It should be possible, for example, to determine the postoperative time required for substances from the graft to reach the homologous organ of the host, together with the subsequent rate of transfer. These data, correlated with the increase in mitotic rate and rate of increase in weight of the host organ, should permit a critical evaluation of the "building-block" hypothesis suggested here, as compared with a possible template or catalytic mechanism. The findings in the investigation thus far strongly suggest that a "building-block" mechanism is responsible, but they do not rule out completely the involvement of a catalytic mechanism. The hypotheses are not mutually exclusive. A second question of interest concerns the intracellular site of localization of the transferred molecules, particularly as related to present concepts of the site of protein synthesis in the cytoplasm.[34] For example, does the localization of graft substances parallel the intracellular localization of antigens in adult organisms?[35]

In conclusion, let us consider the bearing of the present study on the related problem of the inhibition of tissue differentiation by fragments, extracts, or granular fractions of homologous older organs. It has been reported that extracts of embryonic chick heart inhibit differentiation (pulsation) of growing cardiac tissue in vitro[3] and that the differentiation of embryonic frog tissues (e.g., blood, brain, heart) is suppressed by culturing embryos through the neurula stage in a culture medium containing suitable concentrations of products of the homologous adult tissues.[4] More recently, Shaver has reported that a granular fraction of frog brain arrests brain differentiation.[5] Unfortunately, it is not possible to evaluate these reports critically, since the data have not yet been published. It can be argued on theoretical grounds that the two lines of evidence involving stimulation of growth and suppression of differentiation are not in conflict. It is clear in the case of both spleen and kidney grafts studied in this laboratory that the pattern of histogenesis of the host organs is unaltered. One might ask, however, whether a splenic graft made to the chick embryo early in the fourth day of development, well before the spleen has made its appearance, might affect differentiation rather than growth. All too frequently, such apparent conflicts are written off as being due simply to the antagonistic nature of growth and differentiation, a most uninstructive conclusion. A detailed study of the differences in methodology, conditions of host or affected embryos (normal, aged, or overripe, etc.) and state of development of target tissues may provide a basis for interpreting the present differences in results and, in addition, afford another point of departure in our analysis of the mechanisms of growth and differentiation.

* A joint contribution from the Pathology Division, Medical Department, Brookhaven National Laboratory (sponsored by the Atomic Energy Commission), and the Department of Zoölogy, Indiana University (Contribution No. 550), supported by grants from the American Heart Asso-

ciation and the National Science Foundation. It is a pleasure to acknowledge the advice and co-operation of the staff members of the Pathology and Biochemistry divisions, Brookhaven National Laboratory, and the expert technical assistance of J. F. Albright and A. M. Mun.

[1] P. Weiss and H. Wang, *Anat. Rec.*, **79**, 62 (1941); P. Weiss, *Yale J. Biol. Med.*, **19**, 235–278 (1947); P. Weiss, *Quart. Rev. Biol.*, **25**, 177–198 (1950).

[2] P. Weiss, *Arch. Neerland. Zool.*, **10** (Suppl. 1), 159–161 (1953).

[3] P. Weiss, *Science*, **115**, 487–488 (1952).

[4] S. M. Rose, *Anat. Rec.*, **113**, 527 (1952).

[5] J. Shaver, paper presented before American Society of Zoölogists, December 29, 1953.

[6] V. Danchakoff, *Am. J. Anat.*, **24**, 127–189 (1918); C. M. Pomerat, *Exp. Cell Res.*, Suppl. 1, pp. 578–581 (1949).

[7] B. H. Willier, *Am. J. Anat.*, **33**, 67–103 (1924).

[8] J. D. Ebert, *Physiol. Zool.*, **24**, 20–41 (1951).

[9] J. D. Ebert, *Ann. New York Acad. Sci.*, **55**, 67–84 (1952).

[10] J. D. Ebert, *Anat. Rec.*, **117**, 523 (1953).

[11] D. Brüel, H. Holter, K. Linderstrøm-Lang, and K. Rozits, *Biochim. biophys. acta*, **1**, 101–125 (1947).

[12] J. B. Niederl and B. Niederl, *Micromethods of Quantitative Organic Analysis* (2d ed.; New York: John Wiley & Sons, Inc., 1942).

[13] J. D. Ebert and A. M. Mun, unpublished data.

[14] L. L. Forker, I. L. Chaikoff, C. Entenman, and H. Tarver, *J. Biol. Chem.*, **188**, 31–35 (1951).

[15] D. D. Van Slyke, R. Steele, and J. Plazin, *J. Biol. Chem.*, **192**, 769–805 (1951).

[16] W. Bernstein and R. Ballentine, *Rev. Scient. Instr.*, **20**, 347–349 (1949).

[17] W. C. Schneider, *J. Biol. Chem.*, **161**, 293–303 (1945).

[18] Z. Dische, *Microchemie*, **8**, 4–32 (1936).

[19] John Bonner, *Morphogenesis* (Princeton, N.J.: Princeton University Press, 1952).

[20] J. D. Ebert, *Proc. Indiana Acad. Sci.*, **62**, 72–80 (1953); J. D. Ebert, these PROCEEDINGS, **39**, 333–344 (1953); S. M. Rose, *Am. Naturalist*, **86**, 337–354 (1952).

[21] P. Weiss, in *The Chemistry and Physiology of Growth*, ed. A. K. Parpart, pp. 135–186 (Princeton, N.J.: Princeton University Press, 1949).

[22] D. Rittenberg, E. E. Sproul, and D. Shemin, *Federation Proc.*, **7**, 180 (1948).

[23] F. Friedberg, N. P. Schulman, and D. M. Greenberg, *J. Biol. Chem.*, **173**, 437–438 (1948).

[24] F. Lipmann, *Federation Proc.*, **8**, 597–602 (1949).

[25] F. Haurowitz, *Chemistry and Biology of Proteins* (New York: Academic Press, 1950).

[26] J. H. Northrop, in *The Chemistry and Physiology of Growth*, ed. A. K. Parpart, pp. 3–48 (Princeton, N.J.: Princeton University Press, 1949).

[27] H. Borsook, C. L. Deasy, A. J. Haagen-Smit, G. Keighley, and P. H. Lowy, *J. Biol. Chem.*, **179**, 705–719 (1949).

[28] C. B. Anfinsen and D. Steinberg, *J. Biol. Chem.*, **189**, 739–744 (1951); D. Steinberg and C. B. Anfinsen, *J. Biol. Chem.*, **199**, 25–42 (1952).

[29] T. Peters, Jr., *J. Biol. Chem.*, **200**, 461–470 (1953).

[30] O. Snellman and C. E. Danielsson, *Exp. Cell Res.*, **5**, 436–442 (1953).

[31] M. D. Francis and T. Winnick, *J. Biol. Chem.*, **202**, 273–289 (1953).

[32] H. Muir, A. Neuberger, and J. Perrone, *Biochem. J.*, **49** (suppl.), 55 (1951).

[33] H. O. Halvorson and S. Spiegelman, *J. Bact.*, **64**, 207–221 (1952).

[34] J. Brachet, *Symp. Soc. Exp. Biol.*, **6**, 173–200 (1952).

[35] F. Haurowitz, C. F. Crampton, and H. H. Reller, *Arch. exper. Path. Pharmakol.*, **219**, 11–16 (1953).

NEURAL BASIS OF DIURNAL PERIODICITY IN RELEASE OF OVULATION-INDUCING HORMONE IN FOWL

By R. M. Fraps

AGRICULTURAL RESEARCH SERVICE, UNITED STATES DEPARTMENT OF AGRICULTURE, BELTSVILLE, MARYLAND

Communicated by B. H. Willier, March 11, 1954

Everett[1] has recently reviewed the extensive evidence favoring the hypothesis that spontaneous and reflex-induced ovulation are both controlled by similar hypothalamico-hypophyseal mechanisms. This mechanism, in the rat, is believed to "be activated by a hypothalamic center in which resides the property of spontaneity and which undergoes a well-defined diurnal fluctuation of excitability." The potentiation of excitability in the rat by estrogen and progesterone is emphasized.

Like the rat, the hen ovulates "spontaneously," and the release of ovulation-inducing hormone (OIH ≡ the luteinizing hormone, LH) appears to be controlled by a nervous mechanism. The neural apparatus differs however, in at least some respects, from that of the rat. Certain barbiturates which block ovulation in the rat[2] induce premature ovulation in the hen.[3] Two of these, diallylbarbituric acid (Dial) and pentobarbital sodium (Nembutal), enhance in the hen the ovulation-inducing action of progesterone administered at near subovulatory levels. In contrast, phenobarbital sodium does not cause premature ovulation in the hen,[3] but, like Nembutal in the rat,[4] it blocks progesterone-induced ovulation when the steroid is injected at usually effective levels.[3, 5] In this connection it is significant that progesterone is without direct effect on the ovarian follicle of the hen,[6] and it does not induce ovulation more effectively when placed on the exposed pituitary than when applied at distant sites.[7]

The diversity of the hen's ovulatory responses to barbiturates in itself constitutes good evidence that the OIH release mechanism includes a neural component.[8] In addition, Zarrow and Bastian[9] have recently reported both adrenolytic and parasympatholytic drugs to block naturally occurring as well as progesterone-induced ovulation. These authors suggest that a neural link, possibly similar to that postulated for the rat and rabbit, must be present in the mechanism regulating ovulation in the hen. It may be noted also that Huston and Nalbandov[10] and van Tienhoven[11] have reported experimentally induced irritation in certain regions of the hen's oviduct to suppress ovulation without follicular deterioration, an effect assumed to be mediated by neurogenic suppression of the episodic release of OIH (or LH) required for ovulation.[10–12]

While nervous control of OIH release in the hen thus seems fairly certain, little direct evidence has been adduced for the "well-defined diurnal fluctuation in excitability" so convincingly demonstrated by Everett and Sawyer in the rat.[2] In the ringdove, Dunham and Riddle[13] found estradiol benzoate to cause a 24-hour delay in ovulation of the second egg in the typical 2-egg clutch of this species. Subsequently, it was reported from this laboratory that estradiol benzoate suppressed ovulation (or OIH release) in the hen under conditions believed to indicate diurnal periodicity in sensitivity of the postulated neural apparatus.[8] The results of a more extended study of estradiol benzoate–suppressed ovulation in the hen,

presented here, suggest that estrogen decreases excitability of the neural component of the OIH release mechanism during the hours of normally high excitability. However, the action of estradiol on first ovulations of the cycle is different from its action on second (or subsequent) ovulations. These differences are explained in terms of a hypothesis formulated to account for expressions of diurnal periodicity seen in timing of the succession of ovulations constituting the hen's normal ovulation cycle. Because of the significance, experimentally and theoretically, of time relationships appearing in the ovulation cycle, these are described briefly at the outset.

THE OVULATION CYCLE

In the regularly laying hen the time of occurrence of any ovulation can be estimated fairly satisfactorily from records of lay.[14] Under usual conditions of lighting and maintenance, ovulation takes place only within some 8 hours of the 24. In hens of interest here, ovulation typically occurs on each of two or more consecutive days, fails to occur on one day, and then takes place on the following day to initiate another series. The series of ovulations occurring on consecutive days is known as a sequence. The sequence, together with the single day of "missed" ovulation, constitutes the ovulation cycle. In the restricted sense in which the term is used throughout this paper, the cycle is understood to include events from the first, or C_1, ovulation of one sequence to, but not including, the C_1 ovulation of the succeeding sequence. Sequence length is the number of ovulations occurring on successive days; cycle length refers to the number of days elapsing from the C_1 ovulation of one cycle to the C_1 ovulation of the succeeding cycle and is numerically the same as sequence length plus 1.

Except under limiting conditions in long cycles, each ovulation beyond the first of the cycle occurs later in the day than did its predecessor.[14] This difference in *times of day* at which ovulations take place on successive days is denoted lag and corresponds to the "asynchronous rhythm" of Bastian and Zarrow.[15] Lag decreases, in general, with increasing cycle length. The greatest value of lag is found in the 2-member sequence (3-day cycle), where it is commonly of the order of 4–5 hours. In long cycles, lag is typically greatest at the second ovulation (2.5 to 3.5 hours), decreasing thereafter to a minimal value which may approximate or equal zero hours, and then increasing with the last ovulation or two in the cycle.

Since successive ovulations within a cycle occur at later hours on consecutive days, the first, or C_1, ovulation of the next cycle occurs at an *earlier* hour on the second day following the last, or C_l, ovulation of the preceding cycle. The interval between C_l and C_1 ovulations varies, therefore, from about 40 hours in long cycles to as much as 44 hours in a 3-day cycle. The terminal 16 hours of these intervals between C_l and C_1 ovulations are those of the 24 within which ovulation does not normally take place. The same hours recur each day of the cycle but are of special significance in the interval between C_l and C_1 ovulations. The follicle passing through the 16 terminal hours of the last day of the cycle is highly responsive, and the anterior pituitary body is capable of secreting OIH to effect its ovulation, at any time during these hours.[6, 14] Failure of ovulation to occur during the terminal 16 hours of the last day of the cycle consequently rests in failure of the OIH "release-provoking system" during a similarly circumscribed part of the 24-hour day.[6] The existence of diurnal periodicity in some component of the OIH

release mechanism is therefore clearly expressed in the timing of events in the ovulation cycle.

MATERIALS AND METHODS

The hens used included White Leghorns, Rhode Island Reds, and crosses between these breeds. The birds were individually caged in laying batteries, under electric lights from 6:00 A.M. through 8:00 P.M., and had access to feed and water at all times. The usual records of lay at hourly intervals from 8:00 A.M. through 4:00 P.M. served as the basis of selection of experimental hens. Most of the hens were laying in 2-egg sequences, a few in longer sequences, but all were selected on the expectation that only a single day would elapse between lay of the terminal egg of one sequence and the first egg of the following sequence. Neglecting the few hens laying in sequences of more than 2 eggs, the first, or C_1, ovulation was estimated[14] to occur at about 6:00 A.M., the second, or C_2, ovulation at about 10:30 A.M. of the following day. The interval between C_1 and C_2 ovulations was thus approximately 28.5 hours. Following C_2 ovulation at 10:30 A.M., the C_1 ovulation of the next sequence takes place at about 6:00 A.M. of the second day thereafter, that is, after an interval of 43.5 hours. On the basis of these relationships, estradiol benzoate was injected 14–41 hours before expected C_1 ovulation, 11–28 hours before expected C_2 ovulation. The estrogen, dissolved in corn oil (Mazola) in concentrations of 5 mg. per milliliter, was found to be about equally effective following intramuscular injection at levels of 1–2.5 mg. per hen. The higher level was used in the great majority of experiments.

The occurrence of normal or suppressed ovulation was determined by palpation, according to well-established procedures,[3] for presence or absence of the oviducal egg. When ovulation was suppressed, palpation was timed, through the second day particularly, to give approximate information on the hour of the delayed ovulation. The reliability of palpation in diagnosis of results was established by autopsy of a group of nine hens at 12–14 hours following the occurrence of C_2 ovulations delayed for one day, or at about 60 hours following the administration of estradiol benzoate. The ovaries of these hens were normal in all respects. The absence of atresia in any maturing follicle was conclusive evidence that the estrogen was not causing selective atresia (of the next ovulable follicle) when delay in ovulation was indicated by palpation.

One small but prevading element of error in results arises out of the failure of an occasional hen, however regular her previous record, to ovulate as anticipated. Of any considerable number of hens selected for C_1 ovulation some 5–10 per cent may fail to ovulate as expected. Failure of C_2 ovulation to occur as anticipated is less frequent. Ordinarily such failures are of little consequence, but obviously they must be borne in mind in evaluating any very low incidence of presumed delays in ovulation.

RESULTS

Under "Ovulations Suppressed," in Table 1, are recorded the numbers and percentages of injected hens failing to ovulate on the day of expected normal ovulation, regardless of duration of delay. Numbers and percentages of delayed ovulations (1 day or more than 1 day) are also given in terms of all injected hens in respective

groups. Ovulations delayed by more than 1 day usually occurred on the second day following initial suppression, but a few failed to occur until the third day.

As is evident in Table 1, estradiol benzoate suppressed few expected C_1 ovulations; the incidence of actually suppressed ovulations is almost certainly even less than indicated, since, as has been noted, a small proportion of anticipated C_1 ovulations may fail to occur on the day following completion of a sequence. Taking the results of Table 1 at face value, however, the effect of estradiol either in suppressing C_1 ovulation or in delaying it by 1 day is fairly independent of the time elapsing from injection to normally expected ovulation. Of the total of 12 suppressed ovulations, 8, or 67 per cent, were delayed by only a single day. All delayed C_1 ovulations subsequently occurred at or close to the hour of normally expected C_1 ovulation, therefore at 24 hours, or some multiple of 24 hours, beyond the hour of normally expected ovulation.

In contrast with its extremely limited effect on C_1 ovulation, estradiol benzoate suppressed C_2 ovulation in 66, or 55 per cent, of the total of 119 treated hens (Table 1). Of the 66 hens in which C_2 ovulation was suppressed, 54, or 82 per cent,

TABLE 1

INCIDENCE OF SUPPRESSED AND DELAYED OVULATIONS FOLLOWING INJECTION OF ESTRADIOL BENZOATE AT INDICATED HOURS BEFORE EXPECTED NORMAL OVULATION

Expected Normal Ovulation	Injection (Hours before Expected Ovulation)	No. Hens Injected	Ovulations Suppressed		Ovulations Delayed			
					1 Day		>1 Day	
			No.	Per Cent	No.	Per Cent	No.	Per Cent
C_1 at	14	56	7	13	5	9	2	4
6:00	24	16	3	19	2	13	1	6
A.M.	41	16	2	13	1	6	1	6
	14–41	88	12	14	8	9	4	5
C_2 at	11	13	9	69	6	46	3	23
10:30	15	15	7	47	5	33	2	14
A.M.	19	60	31	52	28	47	3	5
	28	31	19	61	15	48	4	13
	11–28	119	66	55	54	45	12	10

ovulated on the day thereafter. The incidence either of suppressed C_2 ovulations, or of C_2 ovulations delayed by a single day, was not related to the increasing interval from injection to normally expected ovulation (Table 1); in other words, increasing the time over which estradiol benzoate might act to suppress the incidence of C_2 ovulations did not increase its effectiveness.

Delayed C_2 ovulations always occurred within the restricted hours of the 24 during which ovulation normally takes place, that is, from approximately 6:00 A.M. to 2:00 P.M. Most of the delayed ovulations occurred, however, at around 6:00 A.M., indicating that when ovulation of the C_2 follicle is delayed, this follicle subsequently behaves, as a rule, like a normal C_1 follicle. When C_2 ovulation is delayed by 1 day, the delay is therefore actually of the order of 19–20 hours (10:30 A.M. of one day to 6:00 A.M. of the next), not 24 hours, as is usually the case when the C_1 ovulation of the hen, or the second of the clutch in ringdoves, is suppressed.

Estradiol dipropionate and diethylstilbestrol have much the same suppressing action on ovulation as does estradiol benzoate. The differential effect on C_1 and C_2 ovulations is of approximately the same order as that found with estradiol ben-

zoate following injection of 2–5 mg. per hen some 14 hours before expected C_1 ovulation, or 19 hours before expected C_2 ovulation. It is of some interest that injection of as much as 5 mg. (per hen) of estradiol benzoate or dipropionate, or of diethylstilbestrol, did not appreciably increase the incidence of suppressed C_1 or C_2 ovulations.

DIURNAL PERIODICITY IN THE EXCITATION CYCLE

Since the incidence of suppressed ovulations (C_1 or C_2) was found to be independent of the time over which estradiol might act, it is not likely that estrogen interfered to any great extent either with processes of follicular maturation or with the capacity of the anterior pituitary body to produce and to secrete gonadotrophin for ovulation. On these grounds alone, it seems probable that the suppression of ovulation observed in the hen might be attributed to decreased excitability of the postulated neural component of the OIH release mechanism during the relatively limited hours of normally high excitability. The inference of diurnal periodicity in the neural apparatus is corroborated in the observation that ovulation, following suppression, always occurred during those hours of the 24 within which the event normally takes place. However, the full import of the ovulatory responses to estrogen can be evaluated critically only on the basis of some consideration of the possible role of diurnal periodicity in the normal cycle.

1. *The Normal Excitation Cycle.*—Excitation refers here to onset of the specific neural activity which results, over whatsoever pathways, in the episodic release of OIH from the anterior pituitary body, and consequently in ovulation. In cycles of given length, each excitation is assumed to precede the corresponding ovulation by a constant interval. The excitation-ovulation interval may be taken as of the order of 8 hours, a figure based on time elapsing from injection of progesterone to subsequent ovulation of the C_1 follicle.[16] If the excitation-ovulation interval is constant (regardless of its absolute value), successive excitations in excitation cycles must stand in the same relationship to each other as do successive ovulations in ovulation cycles.

The neural apparatus is assumed to respond to definite levels of hormones from the ovary, and, more specifically, from the maturing follicle; these are denoted excitation hormones. Some assumption must obviously be made as to the course by which blood levels of excitation hormones increase with time. Perhaps the simplest view is to suppose excitation hormone levels associated with follicles maturing on successive days to be substantially the same at equal times following the preceding excitation (or OIH release or ovulation). This means that each curve of increasing hormone concentrations, beyond the first of the follicular sequence, is displaced, *in time of day*, by the extent of lag at the preceding excitation.

The relationships implicit in the foregoing premises are represented formally for a 7-day cycle in Figure 1. Shaded abscissae indicate the hours of darkness, 8:00 P.M.–6:00 A.M. Hours of the cycle day are 0–24. The curve passing through E_1, E_2, . . ., E_6 and E_1' describes the postulated diurnal rhythm in neural thresholds; here thresholds (of response) represent the inverse of excitability or sensitivity in the neural apparatus. This curve is understood to be repeated each day of the cycle. Ordinates to the curve measure levels of excitation hormones required for neurogenic response. Curves C_1, C_2, . . ., C_6 and C_1' represent terminal concentra-

tions of excitation hormones associated with follicles of the same designations. Curves C_2, C_3, . . ., C_6 reach values required for excitations E_2, E_3, . . ., E_6 *at later hours on successive days.* Curve C_1', however, though displaced only by the extent of lag at E_6, fails to attain to threshold values on day 7 of the cycle but does so at zero hour of the following day to initiate a new cycle. The first excitation, E_1, of the complete cycle shown in Figure 1 is assumed to have occurred under the same conditions as did E_1'.

The characteristics of timing in ovulation cycles follow clearly from the relationships shown in Figure 1. The restriction of ovulation to within about 8 hours of the 24-hour day—the most obvious aspect of diurnal periodicity—depends upon the appearance of relatively low thresholds in the neural apparatus (i.e., high sensitivity or excitability) during a similarly restricted part of the 24 hours. The characteristic pattern of lag is seen to arise out of an orderly change in relationship between successively displaced excitation curves and variation in neural thresholds

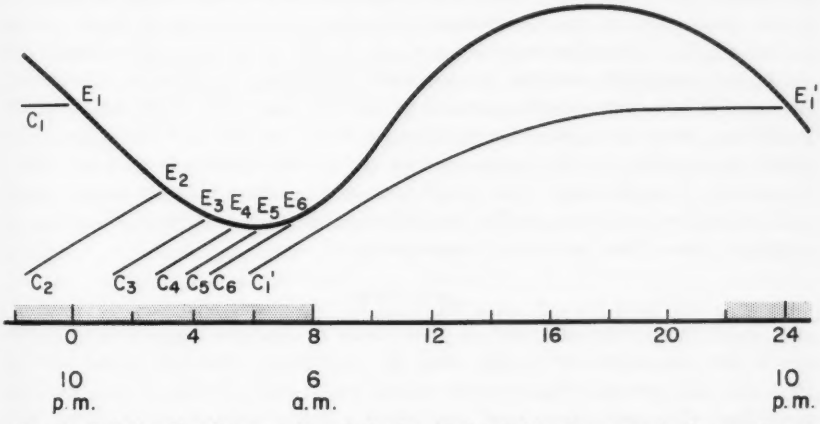

FIG. 1—Relationships between diurnal rhythm in thresholds of response in the neural apparatus and excitation hormone concentrations in a 6-member sequence followed by 1 day with no excitation, thus constituting a 7-day cycle.

with time of day. Failure of the C_1 follicle to ovulate on the last day of the cycle results from the *relatively* high thresholds exhibited by the neural apparatus during some 16 hours of every day, but evident only when the excitation hormone curve associated with a given follicle is so displaced that excitation cannot occur during these hours. But if excitation hormone concentrations associated with such a follicle, the C_1' of Figure 1, continue to increase during these terminal 16 hours of the last day of the cycle, excitation of the neural apparatus should occur finally at higher thresholds than will be possible elsewhere in the cycle. In the hen, this means at an earlier time of day.

2. *The Estrogen-modified Cycle.*—Neglecting small differences associated with cycle length, the first excitation of the 3-day cycle may be supposed to occur at zero hour in Figure 1, the second (and last) at about hour 4.5, thus at about the time of E_3 in the figure. Excitation hormone curves C_1, C_3, and C_1' of the figure may likewise be supposed to represent curves C_1, C_2, and C_1' in the 3-day cycle.

It is clear that even a moderate increase in neural thresholds during hours when such thresholds are normally low would result in failure of excitation to occur between, for example, hours 4 and 8 of the cycle day. Suppression of the expected C_2 ovulation would, of course, follow. A considerably greater increase in thresholds (that is, decrease in excitability) would be required to prevent excitation E_1, normally occurring at zero hour of the cycle day. The results given in Table 1, showing the much lower incidence of suppressed C_1 than of C_2 ovulations, are thus in good accord with theoretical expectations.

It has already been pointed out that a large proportion of suppressed C_2 ovulations (82 per cent) occurred following delay of a single day. Estradiol benzoate is known to be effective for several days, and any increase in neural thresholds caused by this estrogen should be as evident on the day following suppression of ovulation as on the previous day. If, however, the C_2 follicle continued to produce excitation hormones at about the usual rate following failure of excitation to occur at the normal time, concentrations of these hormones on the day thereafter should be of the same order as are concentrations normally associated with the C_1 follicle (C_1' of Fig. 1). Thus the incidence of C_2 ovulations delayed for more than 1 day should be comparable with the incidence of C_1 ovulations delayed by a single day. An examination of the results recorded in Table 1 shows that, of 66 suppressed C_2 ovulations, only 12 or 18 per cent, failed to ovulate on the day thereafter. This figure, 18 per cent, is to be compared with the *initial* suppression of 14 per cent of 88 possible C_1 ovulations. The 19–20-hour delay (rather than 24 hours) typical of C_2 ovulations occurring on the day following their suppression is also to be expected in view of the foregoing considerations.

DISCUSSION

A most interesting outcome of this study of estrogen-suppressed ovulation lies in the consonance of results with the hypothesis proposed to account for timing in the normal ovulation cycle of the hen. Nevertheless, it is to be recognized that "the excitation cycle" represents a highly formal and no doubt oversimplified statement of relationships much in need of further analysis. The conveniently designated excitation hormones, for example, may include effects of estrogen, probably of progesterone or of a physiologically equivalent steroid,[17] and possibly of androgen.[18] The diurnal rhythm in neural thresholds may reside in a well-defined hypothalamic center, but thresholds in this postulated center would almost certainly be subject to neural stimuli arising elsewhere,[19] and these, in turn, to possible modification by one or more of the several "excitation" hormones. The significance of the "neural connection" between oviduct and pituitary gland[10] and that of the theory of the ovulation cycle developed by Nalbandov[12] on the basis of this relationship also calls for attention. Despite the claims of these and of other questions of specific mechanisms involved, the proposed hypothesis of the excitation cycle may serve to emphasize the fundamental importance of diurnal periodicity in the interpretation both of normal cycle relationships and of a considerable body of experimental results.

In contrast with the suppression of ovulation by estrogen in the hen, appropriately timed injection of estrogen advances ovulation in the 5-day cycling rat by about 24 hours.[20] Advancement of ovulation in this species is attributed to *in-*

creased excitability of the presumably hypothalamic center in consequence of increased estrogen levels;[20, 21] suppression of ovulation in the hen has been attributed to *decreased* excitability of a comparable apparatus. We are confronted here with another of those dissimilar responses in the rat and in the hen for which no satisfactory explanation can be offered at present. A clue to these differences may possibly lie in Harris' observation that the hypothalamus may be "regarded as a region where the effect of neural stimuli, arising in the external environment and in the higher centers of the brain, are integrated with humoral factors."[19] Conceivably, these "stimuli" from higher centers may include inhibiting as well as stimulating influences. Quantitative differentials in the effects of certain agents (e.g., estradiol benzoate) on neural centers outside the hypothalamic center immediately concerned in activation of the hypophysis might then lead to completely contrary responses such as advanced or delayed ovulation, in the rat and the hen. The differing and sometimes opposed effects of different barbiturates on ovulation in the hen may have a similar basis in differential neural responses antecedent to the final reaction of the presumed hypothalamic center to excitation hormones.

A direct comparison of the effect of estradiol in the hen and the ringdove is not possible, since Dunham and Riddle were able to observe delayed ovulation only of the second egg of the ringdove's clutch.[13] Moreover, ovulation of the first egg of the ringdove's clutch takes place around midnight, while the second ovulation occurs 4–6 hours *earlier* on the second day thereafter. Timing of first and second ovulations of the dove's clutch is thus not comparable with timing of C_1 and C_t ovulations of the hen's sequence, but rather with the C_t ovulation of one sequence and the C_t ovulation of the succeeding sequence. In both species, it is the ovulation normally occurring earliest, *in time of day*, that is delayed by 24 hours or some multiple thereof. The existence of diurnal periodicity in some element of the ovulation-inducing mechanism of the dove seems implicit in the 24-hour delay, as was recognized by Dunham and Riddle. It seems likely that a neural apparatus similar to that believed to exist in the hen is involved also in the periodicity observed in the ringdove. If so, the "lateness" of the first ovulation in the dove may result from partial suppression of neural thresholds in consequence of higher estrogen levels presumably associated with ovulation of the first follicle of the pair.[13]

SUMMARY

Estradiol benzoate injected intramuscularly into regularly ovulating hens at levels of 1.0–2.5 mg. per hen suppressed 14 per cent of expected first (C_1) ovulations of the cycle, in contrast with 55 per cent of expected second (C_2) ovulations in similar cycles. Sixty-seven per cent of suppressed C_1 ovulations and 82 per cent of suppressed C_2 ovulations occurred within the hours of normally occurring ovulation on the day following suppression. The incidence of suppressed ovulations (C_1 or C_2) and the proportion of ovulations delayed by 1 day were independent, over wide ranges, of the time from injection to normally expected ovulation. These observations constitute strong evidence that the neural mechanism believed to control the release of ovulation-inducing hormone from the anterior pituitary body exhibits diurnal periodicity in its thresholds of response to excitation hormones. This concept is shown to provide a basis for explanation of many of the characteristics of timing of the hen's ovulation cycle, as well as of effects following the administration of estradiol benzoate.

The author wishes to express his appreciation to Dr. M. H. Conner for carrying out many of the experiments. Estradiol benzoate and other estrogens were supplied most generously by the Research Department of Ciba Pharmaceutical Products, Inc., Summit, New Jersey.

[1] J. W. Everett, *Ciba Found. Coll. Endocrinol.*, **4**, 167–178 (1952).
[2] J. W. Everett and C. H. Sawyer, *Endocrinology*, **47**, 198–218 (1950).
[3] R. M. Fraps and J. F. Case, *Proc. Soc. Exp. Biol. Med.*, **82**, 167–171 (1953).
[4] J. W. Everett, *Anat. Rec.*, **109**, 291 (1951).
[5] R. M. Fraps, *Poultry Sci.*, **32**, 899 (1953).
[6] I. Rothchild and R. M. Fraps, *Endocrinology*, **44**, 141–149 (1949).
[7] I. Rothchild, *Federation Proc.*, **8**, 135 (1949).
[8] *Report of Chief of Bureau of Animal Industry* (United States Department of Agriculture, 1953), p. 41.
[9] M. X. Zarrow and J. W. Bastian, *Proc. Soc. Exp. Biol. Med.*, **84**, 457–459 (1953).
[10] T. M. Huston and A. V. Nalbandov, *Endocrinology*, **52**, 149–156 (1953).
[11] A. van Tienhoven, *Anat. Rec.*, **115**, 374 (1953).
[12] A. V. Nalbandov, *Poultry Sci.*, **32**, 88–103 (1953).
[13] H. H. Dunham and O. Riddle, *Physiol. Zoöl.*, **15**, 383–394 (1942).
[14] I. Rothchild and R. M. Fraps, *Endocrinology*, **44**, 134–140 (1949).
[15] J. W. Bastian and M. X. Zarrow, *Proc. Soc. Exp. Biol. Med.*, **79**, 249–252 (1952).
[16] R. M. Fraps and A. Dury, *Proc. Soc. Exp. Biol. Med.*, **52**, 346–349 (1943).
[17] R. M. Fraps, C. W. Hooker, and T. R. Forbes, *Science*, **108**, 86–87 (1948).
[18] A. V. Nalbandov, (unpublished; cited by Huston and Nalbandov, *op. cit.*).
[19] G. W. Harris, *Ciba Found. Coll. Endocrinol.*, **4**, 106–114 (1952).
[20] J. W. Everett, *Endocrinology*, **43**, 389–405 (1948).
[21] C. H. Sawyer, J. W. Everett, and J. E. Markee, *Endocrinology*, **44**, 218–233 (1949).

MENDELIAN AND NON-MENDELIAN INHERITANCE OF STREPTOMYCIN RESISTANCE IN CHLAMYDOMONAS REINHARDI

By Ruth Sager

ROCKEFELLER INSTITUTE FOR MEDICAL RESEARCH, NEW YORK CITY

Communicated by M. Demerec, March 4, 1954

Studies of inheritance in the green alga *Chlamydomonas reinhardi* were initiated in the search for material in which heritable alterations affecting the chloroplast and other cytoplasmic bodies could be analyzed. It has been shown[1, 2] that this organism possesses a readily controlled, complete sexual cycle and that a number of factors affecting mating type,[3] pigment formation,[4] and streptomycin resistance[4] exhibit the expected 2:2 segregation and recombination. In addition, one instance of non-Mendelian segregation has been observed; the character concerned here, too, is streptomycin resistance. It is with the analysis of the nonchromosomal nature of this inheritance that the present report is concerned.

Inheritance of resistance to streptomycin and other growth inhibitors has previously been studied primarily in bacterial systems[5] where conventional methods of genetic analysis were not available.

Materials and Methods.—Nutritional studies with *C. reinhardi*, control of the mating reaction, and methods of genetic analysis are described elsewhere;[2, 4, 6] relevant

points will be summarized here. Vegetative cells are haploid, of either *plus* or *minus* mating type; both multiply indefinitely as vegetative clones. When cells of the two mating types are mixed under suitable conditions,[2] pairing occurs between *plus* (mt^+) and *minus* (mt^-) cells leading to zygote formation. The zygote is diploid and contains, so far as can be detected, all the contents of both parental cells. After a short maturation period, the zygote undergoes meiosis and germinates, with the production of four haploid cells (zoöspores) which show 2:2 segregation for mating type.

The method of crossing consists, in brief, of mass mating of genetically marked strains, isolation and germination of individual zygotes, and recovery from each zygote of four colonies representing the four products of meiosis. Sometimes more than four colonies are recovered, due to a mitotic division following meiosis but before separation of the cells.

The parental cultures used in this investigation, *21gr*, and *4y*, are streptomycin-sensitive, killed by 50 μg of streptomycin per milliliter of medium. Strains *sr-1* and *sr-2* were recovered as resistant colonies from a suspension of ultraviolet-irradiated cells of strain *21gr* spread onto agar containing 100 μg/ml of streptomycin. Subsequently, *sr-1* was found to resist 100 μg/ml and *sr-2* to resist 500 μg/ml streptomycin, as shown by the recovery of equal numbers of colonies on plates with and without streptomycin. These strains have remained resistant after 2 years of repeated subculture in the absence of streptomycin.

TABLE 1

CLASSIFICATION OF PROGENY FROM CROSS OF STRAIN *sr-2* ($sr\ mt^+y_1^+$) × *4Y* ($ss\ mt^-y_1^-$)

	Progeny Classes*	No. of Zygotes
1.	$sr\ mt^+y_1^+$; $sr\ mt^+y_1^+$; $sr\ mt^-y_1^-$; $sr\ mt^-y_1^-$	9
2.	$sr\ mt^+y_1^-$; $sr\ mt^+y_1^-$; $sr\ mt^-y_1^+$; $sr\ mt^-y_1^+$	15
3.	$sr\ mt^+y_1^+$; $sr\ mt^+y_1^-$; $sr\ mt^-y_1^+$; $sr\ mt^-y_1^-$	25
		49

* In four instances only three offspring were recovered per zygote, and the genotype of the fourth clone was inferred.
sr = streptomycin-resistant; *ss* = streptomycin-sensitive.

Since the lethal concentration of streptomycin was found to vary with nutritional conditions, all tests were made under the same conditions (cells grown in light at 25° C. on medium II[7]). The light source was fluorescent tubes at an intensity of 700–800 f-c. Streptomycin sulfate (Lilly) was used in all experiments.

In the crosses to be described between streptomycin-resistant and -sensitive strains, the progeny have been classified for resistance by their ability to grow on agar with and without streptomycin. In sensitive clones, resistant mutants appear so rarely that this method provides a decisive test. In offspring of the original cross, resistance to streptomycin was found to be unimpaired after four subcultures in its absence. Backcross progeny were subcultured only once before testing. Homogeneity of a few selected clones was determined by direct count of colonies recovered on replicate plates with and without streptomycin.

Experimental: Inheritance of sr-1.—Crosses of *sr-1* with sensitive strains will be briefly summarized here and reported in detail elsewhere.[4] The offspring of 35 zygotes have been analyzed, and in each case two of the four offspring clones were found to be resistant to 100 μg/ml streptomycin, and two were sensitive. No link-

age has been found with any other markers studied, but by tetrad analysis it can be shown that the factor behaves as if linked with a centromere and therefore is probably chromosomal.

Inheritance of sr-2.—The inheritance of streptomycin resistance in strain *sr-2* ($mt^+y_1^+$) was investigated by crossing it with strain *4y* ($mt^-y_1^-$); the strains differ in mating type and ability to form chlorophyll in the dark (y_1), as well as in susceptibility to streptomycin. The offspring of 49 zygotes were analyzed (Table 1). All F_1 clones were uniformly streptomycin-resistant upon initial testing and after four subcultures in the absence of streptomycin as well. The other two markers present in the cross, mating type and chlorophyll formation in the dark, segregated 2:2 as in previous studies of these markers.[4]

A further indication that meiosis proceeded normally in this cross is provided by the frequency of recombination between the markers y_1 and *mt*. The ratio found for the three possible tetrad types[8] (parental only; recombinant only; all four types) was 9:15:25 in this cross, a distribution which does not differ significantly from that found for these markers in other crosses.[4]

From these data it is concluded that all classes of progeny have been recovered, that meiosis has proceeded normally, and that other explanations must be sought for the 4:0 segregation of streptomycin resistance.

TABLE 2

INHERITANCE OF STREPTOMYCIN RESISTANCE AMONG PROGENY OF FIRST-GENERATION BACKCROSS

Cross	Phenotype of Progeny Sets: All Resistant (4:0)	All Sensitive (0:4)
913*a* (*sr* $mt^+y_1^+$) × *4y* (*ss* $mt^-y_1^-$)	5	0
913*c* (*sr* $mt^+y_1^-$) × *4y* (*ss* $mt^-y_1^-$)	11	0
913*e* (*sr* $mt^+y_1^+$) × *4y* (*ss* $mt^-y_1^-$)	14	0
913*b* (*sr* $mt^-y_1^-$) × *21gr* (*ss* $mt^+y_1^+$)	0	18
913*f* (*sr* $mt^-y_1^-$) × *21gr* (*ss* $mt^+y_1^+$)	0	11
913*g* (*sr* $mt^-y_1^+$) × *21gr* (*ss* $mt^+y_1^+$)	0	51

Phenotypic Stability of F_1 Clones.—It seemed possible that two of the four F_1 clones recovered from each zygote were only temporarily streptomycin-resistant, as a result of material introduced from the cytoplasm of the zygote. This possibility was tested by growing three F_1 clones from zygote No. 913 [clones 913*a* ($y_1^-mt^+$), 913*b* ($y_1^-mt^-$), and 913*g* ($y_1^+mt^-$)] for about 90 cell generations continuously in the log phase, in an attempt to dilute out any such material. Aliquots of these cultures were plated for colony counts after about 15, 25, 65, 75, and 90 generations. In all cases, approximately equal numbers of colonies appeared on plates with and without 400 μg/ml streptomycin.

On the assumption that the resistance might be due to some cytoplasmic particle dividing at a rate different from that of the cell, attempts were made to lose the resistance phenotype by growing the F_1 clones from zygote No. 913 under a variety of adverse conditions. Neither growth at high temperatures (30° and 35° C.) for 21 cell generations, nor heat shock (1 hour at 46° C.), nor slow growth on a low-nitrogen medium, nor aging in liquid culture for 7, 10, 15, 27, and 34 days after inoculation have led to a loss of streptomycin resistance among the survivors.

No instance has been found of a hereditary loss of resistance by any clone under conditions of vegetative propagation. It has been concluded, therefore, that streptomycin resistance is a permanent hereditary property of the F_1 clones.

Results of Backcrosses.—Further genetic evidence on the nature of the factor conferring streptomycin resistance was obtained by backcrossing the F_1 clones to sensitive strains. Resistant clones of *plus* mating type were backcrossed to *4y*, the strain used in the initial cross; those of *minus* mating type were crossed to *21gr*, the sensitive strain from which *sr-2* was obtained.

The offspring clones of zygote No. 913 were chosen for the first-generation backcross. The results, summarized in Table 2 and diagrammed in Figure 1, indicate that the *plus* clones gave rise to only streptomycin-resistant offspring, as in the case of the parental cross; but the *minus* clones gave rise to only streptomycin-sensitive offspring. In no case were both resistant and sensitive clones found among the offspring of a single zygote. However, as in the parental cross, factors both for mating type and for chlorophyll formation in the dark segregated 2:2, indicating normal meiosis and recovery of all progeny.

A second-generation backcross was carried out with the eight progeny clones of zygote No. 1106, derived from the cross 913*a* × *4y*. As shown in Table 3, the results, in general, are the same as those in Table 2. However, among the 47 zygotes analyzed in crosses of the *minus* clones with *21gr*, five exceptional cases were found in which the progeny were all streptomycin-resistant instead of -sensitive. The significance of these cases will be considered later.

A third-generation backcross was carried out with the progeny of zygote 1325 from the cross 1106*g* × *4y* (Table 3); again the *plus* mating type clones gave rise exclusively to streptomycin-resistant offspring, and the *minus* clones to sensitives, with the two exceptions noted.

Results of Other Crosses.—To determine the generality of observations made on the progeny of zygote No. 913 and its descendants, the four F_1 clones from two other zygotes, Nos. 914 and 916, were also backcrossed to the sensitive strains *21gr* and *4y*. The results, shown in Table 4, are consistent with those previously obtained, showing that the behavior is characteristic of the streptomycin-resistance factor carried by strain *sr-2*.

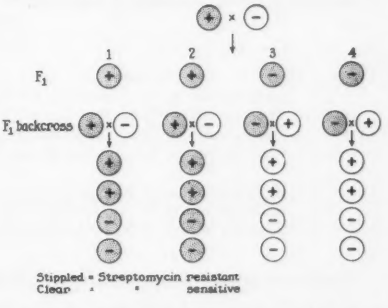

FIG. 1.—Inheritance of streptomycin resistance of strain *sr-2*. *Plus* and *minus* signs refer to mating type. The initial cross, *sr* mt^+ × *ss* mt^-, gave rise exclusively to *sr* offspring, which segregated 2:2 for the markers *mt* and y_1. F_1 clones of *plus* mating type backcrossed to *ss* mt^- produced all *sr* offspring (4:0), but F_1's of *minus* mating type backcrossed to *ss* mt^+ produced only sensitive progeny (0:4).

Crosses among resistant F_1 clones have resulted entirely in resistant offspring in 99 progeny sets tested. When one set of these offspring were backcrossed to sensitive strains, the *plus* mating types again gave rise to resistant offspring, and the *minus* mating type clones to sensitives.

To test the effect of other strains of *plus* mating type on transmission of resistance, a resistant F_1 clone of *minus* mating type (913*g*) was crossed with two *plus* strains, *sr-1* and a pigment mutant no. 164. Strain *sr-1*, described above, is resistant to 100 μg/ml streptomycin, owing to a single gene difference from the parental strain. In the cross *sr-1* × 913*g*, the F_1 offspring from each zygote segregated 2:2 for resistance to 100 μg/ml and sensitivity. Thus the presence in the zygote of a gene

for resistance to 100 μg/ml from the *sr-1* parent did not prevent the loss of resistance to 500 μg/ml.

Mutant 164 is a brown strain, which is blocked in chlorophyll synthesis, accumulates chlorophyll precursors, and grows rather poorly. Because of evident physiological differences between No. 164 and its normal green parent, it seemed possible that it might give a different result in crosses with 913*g* than did the other streptomycin-sensitive strains. However, all F_1 offspring from the cross 164 $\times$ 913*g* were streptomycin-sensitive, as in all other crosses of 913*g* with sensitive strains.

TABLE 3

INHERITANCE OF STREPTOMYCIN RESISTANCE IN SECOND- AND THIRD-GENERATION BACKCROSSES

BACKCROSS	PHENOTYPE OF PROGENY SETS: Resistant (4:0)	Sensitive (0:4)
Second generation:		
1106*a* (*sr* $mt^-y_1^+$) $\times$ *21gr* (*ss* $mt^+y_1^+$)	1*	13
1106*b* (*sr* $mt^-y_1^+$) $\times$ *21gr* (*ss* $mt^+y_1^+$)	2*	12
1106*c* (*sr* $mt^-y_1^+$) $\times$ *21gr* (*ss* $mt^+y_1^+$)	2*	14
1106*d* (*sr* $mt^-y_1^+$) $\times$ *21gr* (*ss* $mt^+y_1^+$)	0	3
1106*e* (*sr* $mt^+y_1^-$) $\times$ *4y* (*ss* $mt^-y_1^-$)	24	0
1106*f* (*sr* $mt^+y_1^-$) $\times$ *4y* (*ss* $mt^-y_1^-$)	10	0
1106*g* (*sr* $mt^+y_1^-$) $\times$ *4y* (*ss* $mt^-y_1^-$)	8	0
1106*h* (*sr* $mt^+y_1^-$) $\times$ *4y* (*ss* $mt^-y_1^-$)	9	0
Third generation:		
1325*a* (*sr* $mt^-y_1^-$) $\times$ *21gr* (*ss* $mt^+y_1^+$)	0	8
1325*c* (*sr* $mt^-y_1^-$) $\times$ *21gr* (*ss* $mt^+y_1^+$)	2*	14
1325*b* (*sr* $mt^+y_1^-$) $\times$ *4y* (*ss* $mt^-y_1^-$)	21	0

* Exceptional cases in which all offspring were phenotypically resistant instead of sensitive.

TABLE 4

INHERITANCE OF STREPTOMYCIN RESISTANCE AMONG BACKCROSS PROGENY OF ZYGOTE SETS NOS. 914 AND 916

CROSS	PHENOTYPE OF PROGENY SETS: All Resistant (4:0)	All Sensitive (0:4)
914*a* (*sr* $mt^-y_1^-$) $\times$ *21gr* (*ss* $mt^+y_1^+$)	0	9
914*g* (*sr* $mt^-y_1^+$) $\times$ *21gr* (*ss* $mt^+y_1^+$)	0	9
914*c* (*sr* $mt^+y_1^+$) $\times$ *4y* (*ss* $mt^-y_1^-$)	9	0
914*f* (*sr* $mt^+y_1^-$) $\times$ *4y* (*ss* $mt^-y_1^-$)	2	0
916*a* (*sr* $mt^-y_1^-$) $\times$ *21gr* (*ss* $mt^+y_1^+$)	0	6
916*h* (*sr* $mt^-y_1^+$) $\times$ *21gr* (*ss* $mt^+y_1^+$)	0	4
916*d* (*sr* $mt^+y_1^-$) $\times$ *4y* (*ss* $mt^-y_1^-$)	21	0

Backcross of Exceptional Sets.—A few instances have been found, starred in Table 3, in which resistant clones of *minus* mating type, crossed to sensitive, gave all resistant offspring (4:0 segregation) instead of all sensitive offspring (0:4 segregation), as had occurred in most cases. Four sets of such offspring were backcrossed with sensitive strains as in previous crosses. The results, summarized in Table 5, show that here, too, it is the *plus* mating type clones which give rise to resistant offspring and the *minus* clones to sensitives, as diagrammed in Figure 1, despite the fact that in these four sets the resistance was introduced by the *minus* parent. Thus, in general, it is only the progeny clones of *plus* mating type that transmit the resistance to their offspring.

Analysis.—The data which have been presented demonstrate that (1) there is a hereditary factor conferring resistance to 500 μg/ml of streptomycin, which exhibits 4:0 segregation in crosses with sensitive clones of *minus* mating type, and (2) in

backcrosses of resistant progeny with sensitive clones of *plus* mating type the resistance is totally lost, with a few exceptions. What hereditary mechanism is responsible for the 4:0 segregations? How can the 0:4 segregation of half the backcross progeny be explained? For purposes of analysis, the 4:0 and the 0:4 segregations will be considered separately.

1. *The 4:0 Segregation.*—There are two hypotheses which would attribute these data to the behavior of Mendelian genes: (1) the hypothesis of resistance controlled by a single gene linked to mating type, which determines the phenotype of the haploid offspring in the zygote; and (2) a multigene hypothesis.

The first hypothesis appears most unlikely in view of the backcross data of Table 5. In the parental crosses from which these progeny sets were derived, resistance was introduced with the *minus* mating type parent, but in each case it is the *plus* mating type progeny which transmit resistance to their offspring. To account for

TABLE 5

INHERITANCE OF STREPTOMYCIN RESISTANCE AMONG PROGENY OF EXCEPTIONAL CASES*

Cross	Phenotype of Progeny Sets: All Resistant (4:0)	All Sensitive (0:4)
1457*b* (*sr* $mt^+y_1^+$) × *4y* (*ss* $mt^-y_1^-$)	7	0
1457*c* (*sr* $mt^-y_1^-$) × *21gr* (*ss* $mt^+y_1^+$)	0	1
1457*d* (*sr* $mt^-y_1^-$) × *21gr* (*ss* $mt^+y_1^+$)	0	5
1461*a* (*sr* $mt^+y_1^+$) × *4y* (*ss* $mt^-y_1^-$)	14	0
1461*b* (*sr* $mt^+y_1^+$) × *4y* (*ss* $mt^-y_1^-$)	11	0
1461*c* (*sr* $mt^-y_1^-$) × *21gr* (*ss* $mt^+y_1^+$)	0	11
1461*d* (*sr* $mt^-y_1^-$) × *21gr* (*ss* $mt^+y_1^+$)	0	8
1649*a* (*sr* $mt^-y_1^+$) × *21gr* (*ss* $mt^+y_1^+$)	0	17
1649*b* (*sr* $mt^-y_1^+$) × *21gr* (*ss* $mt^+y_1^+$)	0	16
1649*c* (*sr* $mt^+y_1^-$) × *4y* (*ss* $mt^-y_1^-$)	10	0
1649*d* (*sr* $mt^+y_1^-$) × *4y* (*ss* $mt^-y_1^-$)	12	0
1945*a* (*sr* $mt^+y_1^-$) × *4y* (*ss* $mt^-y_1^-$)	8	0
1945*d* (*sr* $mt^+y_1^-$) × *4y* (*ss* $mt^-y_1^-$)	18	0
1945*b* (*sr* $mt^-y_1^+$) × *21gr* (*ss* $mt^+y_1^+$)	0	35
1945*c* (*sr* $mt^-y_1^+$) × *21gr* (*ss* $mt^+y_1^+$)	0	26

* Each of these progeny sets came from a cross of a resistant strain of *minus* mating type with a sensitive of *plus* mating type.

these results in terms of a gene linked to mating type would require the occurrence of a four-strand double crossover in each of these four cases. This is so improbable that the hypothesis has been discarded.

On the second hypothesis, the 4:0 segregation is due to the presence of many unlinked genes, each independently conferring streptomycin resistance. Upon continued backcrossing, such genes, each segregating 2:2, would eventually give rise to a "sensitive" clone. In connection with the 4:0 segregation encountered by Ephrussi and collaborators[9] in the case of "petite colonie" in yeast, L'Héritier developed a set of equations for calculating the probability of recovering a single "recessive" clone in each backcross generation. Applying the same calculations here as in the yeast case, it may be concluded that if the 4:0 segregations are to be attributed to the presence of many genes for resistance, more than 20 genes must have mutated simultaneously in strain *21gr* to produce the one-step mutant, *sr-2*. Since this is improbable, the multigene hypothesis is considered untenable.

It appears, then, that the 4:0 segregation data do not fit the expectation for segregation of a chromosome-carried factor, and other modes of transmission must be

considered. The hypothesis which best fits the data is that resistance is deterance by a nonchromosomal hereditary element introduced into the zygote by the resistant parent and distributed through the zygote to all offspring.

2. *The 0:4 Segregation.*—While the 4:0 segregation can be understood in terms of a nonchromosomal determinant, an additional problem is posed by the total loss of resistance in crosses of resistant clones of *minus* mating type with sensitives of *plus* mating type. Two hypotheses have been considered to account for the 0:4 segregation observed in these crosses. Experiments designed to test them are in progress.

a) One hypothesis assumes that there is "maternal inheritance" in *Chlamydomonas*, i.e., an unequal contribution of cytoplasm by the two parental cells. In that case, cells of *plus* mating type would contribute all or most of the cytoplasm to the zygote, and the *minus* cells little or none. Then the exceptional cases discussed above would be attributed to those instances in which some *minus* mating type cytoplasm had been incorporated into the zygote.

The evidence against this hypothesis is cytological but certainly not decisive. When the fusion process is observed under phase-contrast microscopy, it appears that all the contents of both *plus* and *minus* cells are incorporated into the zygote.

b) Alternatively, it may be assumed that the appearance of "maternal inheritance" is determined by a single modifier gene, either the mating-type locus itself or a locus closely linked to it. On this hypothesis, this modifier gene affects the activity or multiplication of the nonchromosomal element, the *plus* mating type favoring a higher level than the *minus*. The exceptions starred in Table 3 would be due to the presence of enough of the element in occasional resistant cells of *minus* mating type to be transmitted through the zygote. The backcross data of Table 5 are consistent with this view.

The data presented in this paper support the view that the streptomycin-resistance property of strain *sr-2* is inherited in a non-Mendelian manner. The demonstration of extra-chromosomal inheritance in this case, as in others which it resembles (e.g., "petite colonie" in yeast[9] and "poky" in *Neurospora*[10]), depends on genetic evidence. The nature of the physical carrier of this inheritance as well as the mechanism of transmission remains obscure.

Summary.—The inheritance of streptomycin resistance in two mutant strains of *Chlamydomonas reinhardi* has been described. In one strain, *sr-1*, the factor for resistance behaves in segregation and recombination like a Mendelian gene. In the other, *sr-2*, crosses to a sensitive strain produce only resistant offspring. Half these offspring, those of *plus* mating type, when mated with sensitive strains, transmit resistance to all their progeny, but with rare exceptions those of *minus* mating type transmit resistance to none. These data demonstrate that, in this strain, streptomycin resistance is inherited in an extra-chromosomal manner. The mechanism responsible for loss of resistance is as yet unclear.

[1] G. M. Smith (unpublished).

[2] R. Sager and S. Granick, *J. Gen. Physiol.*, in press.

[3] G. M. Smith and D. C. Regnery, these PROCEEDINGS, **36,** 246 (1950).

[4] R. Sager (paper in preparation).

[5] M. Demerec, *J. Bact.*, **56,** 63 (1948); L. L. Cavalli Sforza and J. Lederberg, *Symp. Growth Inhibition and Chemotherapy* (1953).

[6] R. Sager and S. Granick, *Ann. New York Acad. Sci.*, **56**, 831 (1953).

[7] Medium II (pH 6.2) contains, in gm/l: sodium acetate·$3H_2O$, 2 gm.; $NaH_2PO_4 \cdot H_2O$, 3.67 gm.; K_2HPO_4, 1.15 gm.; NH_4NO_3, 0.3 gm.; $MgSO_4 \cdot 7H_2O$, 0.3 gm.; $CaCl_2$, 0.04 gm.; $FeCl_3 \cdot 6H_2O$, 0.01 gm.; sodium citrate·$2H_2O$, 0.5 gm.; trace-metal solution, 10 ml. Trace-metal stock solution contains, in mg/l: H_3BO_3, 100 mg.; $ZnSO_4 \cdot 7H_2O$, 100 mg.; $MnSO_4 \cdot 4H_2O$, 40 mg.; $CoCl_2 \cdot 6H_2O$, 20 mg.; $NaMoO_4 \cdot 2H_2O$, 20 mg.; $CuSO_4$, 4 mg.

[8] K. Mather and G. H. Beale, *J. Genetics*, **43**, 1 (1942).

[9] B. Ephrussi, H. Hottinguer, and J. Tavlitski, *Ann. Inst. Pasteur*, **76**, 419 (1949).

[10] M. B. Mitchell and H. K. Mitchell, these PROCEEDINGS, **38**, 442 (1952).

ERRATA: On Restricted Partitions and a Generalization of the Euler φ Number and the Moebius Function

In the article of the foregoing title appearing in these PROCEEDINGS, **39**, 963–968 (1953), the following corrections should be made:

(1) Relation (8), page 965, should read,

If n is odd,

$$\sum_{d/n} (-1)^d \Phi(k, n/d) = \begin{cases} 0 \text{ if } (k, n) \neq n \\ -n \text{ if } (k, n) = n \end{cases}$$

If n is even,

$$\sum_{d/n} (-1)^d \Phi(k, n/d) = \begin{cases} n \text{ if } k = n/2 \\ 0 \text{ otherwise} \end{cases}$$

(2) Relation (11), page 965, should read,

$$\sum_{d/n} \binom{d-1}{\alpha} (-1)^{d-\alpha-1} \Phi(k, n/d) \equiv 0 \bmod (n),$$

where α and β are the integral solutions of $(n/d)\alpha + \beta = s$ such that $0 \leq \alpha < d$, $0 \leq \beta < n/d$; and $0 \leq s < n$.

C. A. NICOL

INFORMATION TO CONTRIBUTORS

THE PROCEEDINGS is the official organ of the NATIONAL ACADEMY OF SCIENCES and of the NATIONAL RESEARCH COUNCIL for the publication of brief accounts of important current researches of members of the ACADEMY and of the COUNCIL and of other American investigators. The PROCEEDINGS will aim especially to secure prompt publication of original announcements of discoveries and wide circulation of the results of American research among investigators in other countries and in all branches of science.

ARTICLES should be brief. The viewpoint should be comprehensive in giving the relation of the paper to previous publications of the author or of others and in exhibiting, where practicable, the significance of the work for other branches of science. Elaborate technical details of the work and long tables of data should be avoided, but authors should be precise in making clear the new results and should give some record of the methods and data upon which they are based.

MANUSCRIPTS should be prepared with a current number of the PROCEEDINGS as a model in matters of form, and should be typewritten in duplicate with double spacing, the author retaining one copy. Illustrations should be confined to text figures of simple character, though more elaborate illustration may be allowed in special instances to authors willing to pay for their preparation and insertion. Particular attention should be given to arranging tabular matter in a simple and concise manner.

REFERENCES to literature, numbered consecutively, will be placed at the end of the article and short footnotes should be avoided. It is suggested that references to periodicals be furnished in some detail and in general accordance with the standard adopted for the Subject Catalogue of the International Catalog of Scientific Literature, viz., name of author, with initials following (ordinarily omitting title of paper), abbreviated name of journal, with place of publication, series (if any), volume, inclusive pages, year. For example: Montgomery, T. H., *J. Morph.*, Boston, **22**, 731–815 (1911); or, Wheeler, W. M., *Konigsberg, Schr. physik. Ges.*, **25**, 1–142 (1914).

MANUSCRIPTS will be accepted only from members of the Academy or from chairmen of the divisions of the National Research Council who will assume responsibility for the propriety and scientific standards of the paper and should be addressed to *Proceedings of the* NATIONAL ACADEMY OF SCIENCES, Attention Miss Mary D. Alexander, University of Chicago Press, 5750 Ellis Avenue, Chicago 37, Illinois. Articles of which a member of the Academy is author or co-author may extend to 10 printed pages, those by non-members to 5 printed pages; space above these limits will be billed to the author at $21.00 per page. Accompanying the manuscript should be a statement by the author or his representative that he assumes responsibility for the charges if incurred.

GALLEY PROOF will be sent, with the understanding that charges for author's corrections shall be billed to him. Author should therefore make final revisions on the typewritten manuscripts. Page proofs will not be sent.

REPRINTS should be ordered when proof is returned, on the reprint order form which will be inclosed with the proof. The order blank gives a tabular estimation of the cost of reprints; a copy will be supplied at any time upon request.

CONTENTS

Page

BIOCHEMISTRY.—NUCLEIC ACID PRECURSORS AND PROTEIN SYNTHESIS . . . . . *By Arthur B. Pardee* 263

BIOCHEMISTRY.—SYNTHESIS OF AROMATIC COMPOUNDS BY *Neurospora* . . . . . *By E. L. Tatum, S. R. Gross, G. Ehrensvärd, and L. Garnjobst* 271

CHEMISTRY.—EFFECT OF LOCAL-ACTION CURRENTS ON THE IRON POTENTIAL. . . . . *By Herbert H. Uhlig* 276

GEOLOGY.—THE NATURAL C^{14} CONTENTS OF MATERIALS FROM HARD-WATER LAKES . . . *By Edward S. Deevey, Jr., Marsha S. Gross, G. E. Hutchinson, and Henry L. Kraybill* 285

MATHEMATICS.—BOUNDEDNESS AND STATIONARITY OF TIME SERIES . . *By S. Bochner* 289

MATHEMATICS.—ANALYTIC ALMOST-PERIODIC FUNCTIONS. I . . . *By Oscar Goldman* 294

MATHEMATICS.—POLYADIC BOOLEAN ALGEBRAS . . . . . *By Paul R. Halmos* 296

MATHEMATICS.—STUDIES IN THE CONFORMAL MAPPING OF RIEMANN SURFACES. II . . . *By Maurice Heins* 302

MATHEMATICS.—ON THE LIE RING OF A SIMPLE RING . . . . . *By I. N. Herstein* 305

MATHEMATICS.—ON A THEOREM BY MURNAGHAN . . . . . *By E. M. Ibrahim* 306

MATHEMATICS.—ON THE ARITHMETIC NORMALITY OF THE GRASSMANN VARIETY . . . *By Jun-ichi Igusa* 309

MATHEMATICS.—ON KÄHLER VARIETIES OF RESTRICTED TYPE . . . *By K. Kodaira* 313

MATHEMATICS.—ON THE DERIVATION OF THE EQUATIONS OF HYDRODYNAMICS FROM STATISTICAL MECHANICS . . . . . *By Charles B. Morrey, Jr.* 317

MATHEMATICS.—DETERMINATION OF THE PLASTIC YIELD CONDITION AS A VARIATIONAL PROBLEM . . . . . *By T. Y. Thomas* 322

MATHEMATICS.—NOTE ON THE DISTRIBUTION OF ZEROS OF EXTREMAL POLYNOMIALS . . . *By J. L. Walsh and J. P. Evans* 332

ZOOLOGY.—THE EFFECTS OF CHORIOALLANTOIC TRANSPLANTS OF ADULT CHICKEN TISSUES ON HOMOLOGOUS TISSUES OF THE HOST CHICK EMBRYO . . . *By James D. Ebert* 337

ZOOLOGY.—NEURAL BASIS OF DIURNAL PERIODICITY IN RELEASE OF OVULATION-INDUCING HORMONE IN FOWL . . . . . *By R. M. Fraps* 348

GENETICS.—MENDELIAN AND NON-MENDELIAN INHERITANCE OF STREPTOMYCIN RESISTANCE IN *Chlamydomonas reinhardi* . . . . . *By Ruth Sager* 356

ERRATA.—ON RESTRICTED PARTITIONS AND A GENERALIZATION OF THE EULER φ NUMBER AND THE MOEBIUS FUNCTION . . . . . *By C. A. Nicol* 363

age
263
271
276
285
289
294
296
302
305
306
309
313
317
322
332
337
348
356
363